

The  
Indo-European  
Languages

Second Edition

Edited by Mate Kapović

ROUTLEDGE LANGUAGE FAMILY SERIES



---

---

THE

---

---

INDO-EUROPEAN

---

---

LANGUAGES

---

---

*The Indo-European Languages* presents a comprehensive survey of the individual languages and language subgroups within this language family.

With over four hundred languages and dialects and almost three billion native speakers, the Indo-European language family is the largest of the recognized language groups and includes most of the major current languages of Europe, the Iranian plateau and the Indian subcontinent.

Written by an international team of experts, this comprehensive, single-volume tome presents in-depth discussions of the historical development and specialized linguistic features of the Indo-European languages.

This unique resource remains the ideal reference for advanced undergraduate and postgraduate students of Indo-European linguistics and languages, but also for more experienced researchers looking for an up-to-date survey of separate Indo-European branches. It will be of interest to researchers and anyone with an interest in historical linguistics, linguistic anthropology and language development.

**Mate Kapović** is Associate Professor in the Department of Linguistics at the University of Zagreb, Croatia. He currently teaches courses on Indo-European phonology, Indo-European and Balto-Slavic accentuation, Indo-European morphology, and general phonology.

## ROUTLEDGE LANGUAGE FAMILY SERIES

Each volume in this series contains an in-depth account of the members of some of the world's most important language families. Written by experts in each language, these accessible accounts provide detailed linguistic analysis and description. The contents are carefully structured to cover the natural system of classification: phonology, morphology, syntax, lexis, semantics, dialectology, and sociolinguistics.

Every volume contains extensive bibliographies for each language, a detailed index and tables, and maps and examples from the languages to demonstrate the linguistic features being described. The consistent format allows comparative study, not only between the languages in each volume, but also across all the volumes in the series.

- |   |  |
|---|--|
| The Austronesian Languages of<br>Asia and Madagascar<br><i>Edited by Nikolaus Himmelmann &amp;<br/>Sander Adelaar</i> | The Mongolic Languages<br><i>Edited by Juha Janhunen</i>   |
| The Bantu Languages<br><i>Edited by Derek Nurse &amp;<br/>Gérard Philippson</i>                                       | The Munda Languages<br><i>Edited by Gregory D.S. Anderson</i>  |
| The Celtic Languages 2nd Edition<br><i>Edited by Martin J. Ball &amp;<br/>Nicole Müller</i>                           | The Oceanic Languages<br><i>Edited by John Lynch, Malcolm Ross &amp;<br/>Terry Crowley</i>                             |
| The Dravidian Languages<br><i>Edited by Sanford B. Steever</i>  | The Romance Languages<br><i>Edited by Martin Harris &amp;<br/>Nigel Vincent</i>  |
| The Germanic Languages<br><i>Edited by Ekkehard König &amp;<br/>Johan van der Auwera</i>                              | The Semitic Languages<br><i>Edited by Robert Hetzron</i>   |
| The Indo-Aryan Languages<br><i>Edited by George Cardona &amp;<br/>Dhanesh K. Jain</i>                                 | The Sino-Tibetan Languages<br>2nd Edition (forthcoming)<br><i>Edited by Graham Thurgood &amp;<br/>Randy J. Lapolla</i> |
| The Indo-European Languages<br>2nd Edition<br><i>Edited by Mate Kapović</i>   | The Slavonic Languages<br><i>Edited by Bernard Comrie &amp;<br/>Greville G. Corbett</i>                                |
| The Iranian Languages<br><i>Edited by Gernot Windfuhr</i>   | The Tai-Kadai Languages<br><i>Edited by Anthony Diller</i>   |
| The Languages of Japan and Korea<br><i>Edited by Nicolas Tranter</i>  | The Turkic Languages<br><i>Edited by Éva Csató &amp;<br/>Lars Johanson</i>   |
| The Khoesan Languages<br><i>Edited by Rainer Vossen</i>   | The Uralic Languages<br><i>Edited by Daniel Abondolo</i>   |

THE  
INDO-EUROPEAN  
LANGUAGES

Second edition

Edited by  
Mate Kapović

 **Routledge**  
Taylor & Francis Group  
LONDON AND NEW YORK

Second edition published 2017  
by Routledge  
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge  
711 Third Avenue, New York, NY 10017

*Routledge is an imprint of the Taylor & Francis Group, an informa business*

© 2017 selection and editorial matter, editor; individual chapters, the contributors

The right of the editor to be identified as the author of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

*Trademark notice:* Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

This edition adapted from the 1998 Routledge translation of *Le Lingue Indoeuropee* © 1993 Società editrice Il Mulino, Bologna

*British Library Cataloguing-in-Publication Data*

A catalogue record for this book is available from the British Library

*Library of Congress Cataloging-in-Publication Data*

A catalog record for this book has been requested

ISBN: 978-0-415-73062-4 (hbk)

ISBN: 978-1-315-67855-9 (ebk)

Typeset in Times New Roman  
by Apex CoVantage, LLC

---

---

# CONTENTS

---

---

<i>List of illustrations</i>	vii
<i>Preface</i>	xiii
<i>Abbreviations</i>	xv
<b>Indo-European languages – introduction</b> <i>Mate Kapović</i>	1
<b>1 Proto-Indo-European</b>	
<b>1 Proto-Indo-European phonology</b> <i>Mate Kapović</i>	13
<b>2 Proto-Indo-European morphology</b> <i>Mate Kapović</i>	61
<b>3 Proto-Indo-European syntax</b> <i>Thomas Krisch</i>	111
<b>2 Proto-Indo-European and language typology</b> <i>Ranko Matasović</i>	153
<b>3 Anatolian</b> <i>H. Craig Melchert</i>	171
<b>4 Indo-Iranian</b>	
<b>1 Indo-Iranian</b> <i>Leonid Kulikov</i>	205
<b>2 Indo-Aryan</b> <i>Leonid Kulikov</i>	214
<b>3 Iranian</b> <i>Nicholas Sims-Williams</i>	263
<b>5 Greek</b> <i>Rupert Thompson</i>	287
<b>6 Italic</b> <i>Rex Wallace</i>	317
<b>7 Celtic</b> <i>Patrick Sims-Williams</i>	352
<b>8 Germanic</b> <i>Joshua Bousquette and Joseph Salmons</i>	387
<b>9 Armenian</b> <i>Birgit Anette Olsen</i>	421
<b>10 Tocharian</b> <i>Douglas Q. Adams</i>	452

**11 Balto-Slavic**

**1 Balto-Slavic** *Steven Young* 479

**2 Baltic** *Steven Young* 486

**3 Slavic** *Marc L. Greenberg* 519

**12 Albanian** *Alexander Rusakov* 552

*Language Index* 609

*Subject-Author Index* 617

---

---

# ILLUSTRATIONS

---

---

## MAPS

1.1	IE languages today	2
1.2	IE languages in the 1st millennium BCE	6
1.1	The <i>centum-satem</i> branches of IE	28
3.1	Anatolia in the Late Bronze Age	173
3.2	Anatolia in the Iron Age	174
5.1	The Greek dialects	288
6.1	The regions of ancient Italy	318
6.2	The languages of pre-Roman Italy	319
7.1	Latin provincial lapidary inscriptions bearing Celtic personal names	355
7.2	Density of Celtic-looking ancient place-names	356
9.1	Armenian in ancient and medieval times	423
10.1	Kucha (= Kuci) and its dependencies in Han times	453
11.1	Baltic tribes at the beginning of the second millenium ad	487
11.2	Map of Slavic languages	520
12.1	The Albanian dialects	553

## FIGURES

1.2	Sentence scheme 1	116
1.3	Sentence scheme 2	116
1.4	Structure of Homer Iliad 1,8	117
1.5	Structure of Homer Odyssey 4,353	118
1.6	Structure of Homer, Odyssey 20, 48	122
1.7	Structure of Rigveda 1,32,2	133
1.8	Structure of Homer, Iliad 18,468 / 1	134
1.9	Structure of Homer, Iliad 18,468 / 2	135
1.10	Structure of Homer, Iliad 5,97	136
1.11	Structure of Homer, Iliad 5,99	136
1.12	Structure of Homer Iliad 5,100	137
1.13	Structure of Homer, Iliad 5,101	137
1.14	Hettrich 2002: 47 The spectrum of meaning of the Rgvedic instrumental	141
4.1	Chronology of Indo-Aryan languages and texts	215
4.2	Polyglossia in ancient India	216
4.3	Simple vowels	267
4.4	The first and second series of palatals in Iranian	270



5.1	Inherited vowel system of Proto-Attic-Ionic	294
5.2	Vowel system of Proto-Attic-Ionic after compensatory lengthening	295
5.3	Vowel system of classical Attic, showing official spellings after 403 BC	296
6.1	<i>Fibula Praenestina</i> (CIL 1 <sup>2</sup> .3)	320
8.1	Two models of Germanic subgrouping	388

## TABLES

1.1	PIE stops	14
1.2	The reflection of PIE palatovelars in <i>satem</i> languages	23
1.3	Basic reflexes of PIE stops	28
1.4	The reflexes of PIE asyllabic resonants	34
1.5	The reflexes of PIE syllabic resonants	36
1.6	The reflexes of PIE *CR <sub>h</sub> HC	39
1.7	The reflexes of PIE *CR <sub>h</sub> HV	39
1.8	The reflexes of PIE vowels	41
1.9	The reflexes of PIE diphthongs	51
1.10	IE <i>o</i> -stems (singular)	64
1.11	IE <i>o</i> -stems (plural)	64
1.12	IE <i>o</i> -stems (neuter)	65
1.13	IE <i>o</i> -stems (dual)	66
1.14	IE <i>eh</i> <sub>2</sub> -stems (singular)	66
1.15	IE <i>eh</i> <sub>2</sub> -stems (plural)	67
1.16	IE <i>eh</i> <sub>2</sub> -stems (dual)	67
1.17	IE root nouns (singular)	71
1.18	IE root nouns (plural)	71
1.19	IE <i>i</i> -stems (singular)	72
1.20	IE <i>i</i> -stems (plural)	72
1.21	IE <i>u</i> -stems (singular)	73
1.22	IE <i>u</i> -stems (plural)	74
1.23	IE <i>r</i> -stems (singular)	75
1.24	IE <i>r</i> -stems (plural)	75
1.25	IE <i>n</i> -stems (singular)	76
1.26	IE <i>n</i> -stems (plural)	76
1.27	IE <i>s</i> -stems (singular)	78
1.28	IE <i>s</i> -stems (plural)	78
1.29	IE demonstrative pronoun *so – *seh <sub>2</sub> – *tod (singular)	85
1.30	IE demonstrative pronoun *so – *seh <sub>2</sub> – *tod (plural)	86
1.31	IE demonstrative pronoun *so – *seh <sub>2</sub> – *tod (dual)	86
1.32	IE demonstrative pronoun *ey (*is) – *ih <sub>2</sub> – *id (singular)	87
1.33	IE demonstrative pronoun *ey (*is) – *ih <sub>2</sub> – *id (plural)	87
1.34	IE athematic present	93
1.35	IE thematic present	93
1.36	IE present dual endings	94
1.37	IE root-aorist	96
1.38	IE sigmatic aorist	96
1.39	IE thematic aorist	97

1.40	IE perfect	98
1.41	IE middle present	99
1.42	IE middle aorist	100
1.43	IE athematic subjunctive	101
1.44	IE thematic subjunctive	101
1.45	IE athematic optative	102
1.46	IE thematic optative	102
1.47	IE athematic imperative	103
1.48	IE thematic imperative	103
1.49	IE future imperative	103
3.1	Anatolian nominal endings	178
3.2	Anatolian personal pronouns	181
3.3	Anatolian anaphoric pronouns	182
3.4	Anatolian interrogative-relative pronoun	183
3.5	Hittite verbal inflection	185
4.1	Encoding transitivity oppositions: diachronic typological features of some language families	210
4.2	Vedic texts and schools	218
4.3	Abbreviations of texts (text sigla)	220
4.4	Old Indo-Aryan consonantism	222
4.5	Assimilation in consonant clusters before dental stops	230
4.6	The main alternation grades	231
4.7	Old Indo-Aryan case endings	233
4.8	Consonant declension	235
4.9	( <i>A</i> ) <i>nt</i> -declension	235
4.10	<i>a</i> -declension (type <i>devá-</i> )	236
4.11	<i>ī</i> -declension (types <i>devī-</i> and <i>vṛkī-</i> )	237
4.12	PIE sources of case endings	237
4.13	Vedic verbal personal endings	239
4.14	Series of the PIE verbal endings	240
4.15	PIE sources of Vedic personal endings	240
4.16	The Vedic verbal system: examples	243
4.17	The Vedic verbal system: examples	244
4.18	The Vedic system of present stem types	245
4.19	The inventory of the present passive forms attested in the RV and AV	246
4.20	Passive paradigm in early Vedic	247
4.21	Declension of masculine <i>a</i> -stems (PIE stems in * <i>o</i> )	275
4.22	Declension of Av. <i>xratu-</i>	276
4.23	Secondary endings in Avestan	279
4.24	Primary endings in Avestan	280
4.25	Conjugation of the thematic present indicative active	280
4.26	Conjugation of the thematic present indicative middle	280
5.1	The Greek (Ionic) alphabet	289
5.2	Treatment of labiovelars in Attic-Ionic and West Greek	292
5.3	Forms of the first declension ( <i>a</i> -stems) in classical Attic	298
5.4	Forms of the second declension ( <i>o</i> -stems) in classical Attic	299

5.5	Underlying endings of third declension, and consonant stem forms in classical Attic	299
5.6	Forms of the third-declension <i>s</i> -stems, <i>i</i> -stems and <i>u</i> -stems in classical Attic	300
5.7	Underlying forms of the <i>-ēu</i> stems, and the resulting forms in classical Attic	301
5.8	Forms of the personal pronouns	304
5.9	Thematic and athematic primary and secondary endings of the active voice	306
5.10	Endings of the strong and weak aorist	307
5.11	Endings of the middle voice	308
5.12	Forms of the perfect and pluperfect in classical Attic	310
6.1	The phonemes of Classical Latin	325
6.2	Classical Latin consonant-stem endings	330
6.3	Classical Latin <i>o</i> -stem endings	330
6.4	Classical Latin noun paradigms, masculine and feminine genders	331
6.5	Classical Latin noun paradigms, neuter	331
6.6	Classical Latin ‘mixed’ declension	333
6.7	Classical Latin <i>ē</i> -stems	333
6.8	Latin personal pronouns	335
6.9	Classical Latin relative pronoun	336
6.10	Organization of verb system	336
6.11	Latin conjugation classes	337
6.12	Subjunctive mood	338
6.13	Personal endings	339
6.14	Perfect passives	339
7.1	Comparison of the ordinal numerals	353
7.2	The Common Celtic vowel system	362
7.3	Pre-Celtic stops	363
7.4	Common Celtic stops	363
7.5	Brythonic consonantal developments	365
7.6	Gaelic consonantal developments	367
7.7	Nominal stems in <i>-o-</i>	368
7.8	Nominal stems in <i>-ā-</i>	369
7.9	Nominal stem in <i>-n-</i>	369
7.10	A possible origin of Old Irish conjunct endings	373
7.11	A possible origin of Old Irish absolute endings	373
7.12	Old Irish passive forms (present tense)	374
7.13	Old Irish deponent forms (present tense)	374
8.1	Pre-Germanic consonants	391
8.2	IE and Germanic stops	392
8.3	Old High German weak verbs	396
8.4	Germanic strong verbs, classes I–V	398
8.5	<i>a</i> -stem masculine nouns	401
8.6	Examples of other noun classes	402
9.1	The Armenian alphabet	422
10.1	The development of PIE vowels in Tocharian	455
10.2	Tocharian consonants	458

10.3	Tocharian vowels	458
10.4	An example of Tocharian nominal declension	462
10.5	Some examples of Tocharian adjectival declension	464
10.6	Deictic pronouns in Tocharian	465
10.7	First and second person pronouns in Tocharian	466
10.8	Tocharian numbers from ‘one’ to ‘ten’	466
10.9	The declension of ‘one’ in the Tocharian languages	467
10.10	Conspectus of the Tocharian verb	469
11.1	The vowel system of common Baltic	489
11.2	The vowel system of East Baltic	490
11.3	The vowel system of Lithuanian	491
11.4	The vowel system of Latvian	491
11.5	Lithuanian accent paradigms for <i>a</i> -stem nouns	494
11.6	The Proto-Baltic consonant system	496
11.7	Declension of <i>a</i> -stem nouns	502
11.8	Declension of <i>a/ā</i> -stem and <i>u/i</i> -stem simple adjectives in Lithuanian	504
11.9	Declension of <i>a/ā</i> -stem and <i>u/i</i> -stem definite adjectives in Lithuanian	505
11.10	Declension of the Lithuanian third-person personal pronoun <i>jis, ji</i>	507
11.11	Proto-Indo-European to late common Slavic vowels	522
11.12	Late common Slavic vowels	523
11.13	PIE, Baltic, and Slavic accent paradigm correspondences	526
11.14	Accentual paradigms	527
11.15	Proto-Indo-European to late common Slavic consonants	528
11.16	Cluster simplification	530
11.17	First velar palatalization	530
11.18	Second and third velar palatalizations	531
11.19	‘Deiotation’ reflexes	532
11.20	Common Slavic consonant system	533
11.21	Vocalic stems	535
11.22	Consonantal stems	535
11.23	Vocalic stem endings	536
11.24	Consonantal stem endings	536
11.25	OCS 3 pers. pronoun	538
11.26	Interrogative pronouns	538
11.27	OCS adverbs of place, time, manner, direction	539
11.28	Inflection of non-past finite verb forms	540
11.29	Verb classes, following Leskien	540
11.30	Slavic verb form classes	541
11.31	Deverbatives from CSL. * <i>věděti</i> ‘to know’, derivatives of CSL. * <i>ryba</i> ‘fish’	543
12.1	Albanian alphabet	559
12.2	PIE short vowels (in stressed position)	563
12.3	PIE long vowels (in stressed position)	563
12.4	PIE diphthongs (in stressed position)	564
12.5	PIE sonorants	567
12.6	PIE obstruents	569
12.7	PIE * <i>s</i> , * <i>y</i> , * <i>w</i>	570
12.8	Noun declension in singular	574

12.9	Noun declension in plural	575
12.10	Articulated adjectives (unmarked order)	581
12.11	Articulated adjectives (inverted order)	581
12.12	Declension of the personal and demonstrative pronouns	582
12.13	Declension of the possessive pronouns	583
12.14	Cardinal numerals	584
12.15	Verb conjugation (the major conjugational classes)	589
12.16	Conjugation of the verbs <i>jam</i> 'to be' and <i>kam</i> 'to have'	591

---

---

# PREFACE

---

---

This volume is formally the 2nd edition of Ramat and Ramat 1998, itself an English translation of Ramat and Ramat 1993. Although the idea and the scope of the volume have remained much the same, the volume itself has changed a lot (the new editor included), and not only due to the necessary updates, inescapable after almost a quarter of a century has passed since the first edition.

The chapter “Proto-Indo-European: Comparison and Reconstruction” (1st ed.) is now replaced by three completely new chapters (“Proto-Indo-European Phonology,” “Proto-Indo-European Morphology,” and “Proto-Indo-European Syntax”). Instead of the chapter “Sanskrit” (1st ed.), there is a new chapter “Indo-Aryan.” Also, a separate short prechapter “Indo-Iranian” has been added. Instead of the chapters “Latin” and “The Italic Languages” (1st ed.), there is only the new chapter “Italic” now. A new prechapter “Balto-Slavic” has been added to precede “Baltic” and “Slavic.” A new chapter “Fragmentarily Attested Indo-European Languages” was planned, but the text has unfortunately not been finished on time. The chapter “The Indo-Europeans: Origins and Culture” (1st ed.) is completely out. The order of the chapters is new. The only authors remaining from the 1st edition are Nicholas Sims-Williams (“Iranian”) and Patrick Sims-Williams (“Celtic”). Their chapters are updated versions of their chapters in the 1st edition; all the other chapters (written by different authors) are completely new. Some of the authors from the 1st edition have, unfortunately, passed away in the meantime, like the great Calvert Watkins (1933–2013) and Werner Winter (1923–2010). New maps and illustrations were also added.

The idea of the volume is to provide a compact but informative introduction to and overview of Indo-European languages, with the stress on Proto-Indo-European, the oldest well attested IE languages, and the largest IE branches. The aim of the volume is to provide an introduction to beginners or non-Indo-Europeanists, but also to be used as a reference book for scholars already doing research in IE linguistics. However, this is not an introduction to historical linguistics, phonology, morphology, or syntax in general. It is assumed that the potential readers will already be familiar with basic linguistics, although the aim was to make the volume as reader-friendly as possible. The approach in the volume is mostly historical, as traditional in IE linguistics, and the volume is linguistics only – i.e., there are no chapters on IE culture, archeology, mythology, poetics, etc. In the last twenty years, a lot of historical-comparative grammars of IE in English have appeared (cf. the literature in the first three chapters), but usually by one author surveying all the languages – here, specialists in the field deal with one IE branch (or topic) only from a wider IE perspective. The chapters of the volume are meant to be complementary, but each chapter is supposed to be able to stand for itself (obviously, there is some repetition of certain facts).

The editor has tried to make this volume more uniform (e.g., the same PIE “orthography” is used throughout the book) and internally coherent than the first one, but it was impossible to “tame” completely so many authors, each with their own writing style and

varying interests. Obviously, each chapter has the author's personal touch (which is not bad – otherwise, all the introductions to IE linguistics would be the same). Naturally, not all literature and different views that would deserve it could be included in this volume. Also, the reader will sometimes note the conflicting views in the volume – not only was this impossible to avoid, but there was no reason to avoid it. This is something that the reader has to take into account when he or she encounters differing PIE reconstructions or interpretations of certain processes or phenomena. Rather than being a hindrance, this can allow a beginner to understand that historical linguistics is not a dogma, and that by applying its methods we can sometimes get differing results. It is exactly after the competition of different ideas that *communis opinio* is refined. In the end, the editor is glad to note that, although the basic idea of the volume is to provide an introduction and overview of IE linguistics and in spite of the fact that people have been doing IE historical linguistics for quite some time now, the volume does indeed contain quite a few new insights and interpretations, which will make it an interesting read not only for beginners and non-experts but also for professional IE linguists.

Finally, I would like to thank the publisher (especially the editor Samantha Vale Noya) for the opportunity to do this (though I have often, in times of despair, regretted the fact that I accepted the very difficult and often tedious editorial work), and all the authors who have worked hard in order for us to be able to produce such a wonderful volume.

## REFERENCES

- Ramat, Anna Giacalone & Ramat, Paolo (eds.) 1993 [1997], *Le lingue indoeuropee*, II  
Mulino, Bologna
- Ramat, Anna Giacalone & Ramat, Paolo (eds.) 1998, *The Indo-European Languages*,  
Routledge, London–New York

---

# ABBREVIATIONS

---

## LANGUAGE/DIALECT ABBREVIATIONS

AC	Arcado-Cypriot
Aeol.	Aeolic
AI	Attic-Ionic
Alb.	Albanian
Anat.	Anatolian
Arc.	Arcadian
Arm.	Armenian
Att.	Attic
Av.	Avestan
AV	Atharva-Veda
Balt.	Baltic
BCMS	Bosnian/Croatian/Montenegrin/Serbian (= Serbo-Croatian)
Boe.	Boeotian
Bret.	Breton
BRuss.	Belorussian/Belarusian
BSl.	Balto-Slavic
Bulg.	Bulgarian
Čak.	Čakavian
CBalt.	Common Baltic
Celt.	Celtic
Clb.	Celtiberian
Class. Lat.	Classical Latin
CLuw.	Cuneiform Luwian
Corn.	Cornish
Cret.	Cretan
Croat.	Croatian
CS	Church Slavic
CSl.	Common Slavic
Cypr.	Cypriot
Cz.	Czech
Dac.	Dacian
Dan.	Danish
Dor.	Doric
E	East . . .
EBalt.	East Baltic
EGr.	East Greek
ELith.	East Lithuanian
Eng.	English



Est.	Estonian
Fal.	Faliscan
Finn.	Finnish
Fr.	French
Gaul.	Gaulish
GAv.	Gatha-Avestan (= Old Avestan)
Germ.	German
Gl.	Glossing
Gmc	Germanic
Goth.	Gothic
Gr.	Greek
Hitt.	Hittite
HLuw.	Hieroglyphic Luwian
Hom.	Homeric (Homer)
IE	Indo-European
IIr.	Indo-Iranian
Illyr.	Illyrian
Ion.	Ionic
Iran.	Iranian
It.	Italic
Ital.	Italian
Khot.	Khotanese
Lat.	Latin
Latv.	Latvian
LAv.	Later Avestan (= Young Avestan)
LCSl.	Late Church Slavic
Lep.	Lepontic
Lesb.	Lesbian
Lith.	Lithuanian
LSorb.	Lower Sorbian
Luw.	Luwian
Lyc.	Lycian
Lyd.	Lydian
M	Middle . . .
Mac.	Macedonian (Slavic)
ManMPers	Manichean Middle Persian
MBret.	Middle Breton
MDutch	Middle Dutch
Messap.	Messapian
MHG	Middle High German
MIran.	Middle Iranian
MIr.	Middle Irish
MParth	Middle Parthian
MPers.	Middle Persian
MW	Middle Welsh
Myc.	Mycenaean
N	North . . .
NE	North-East
Norw.	Norwegian

---

NPhryg.	New Phrygian
NW	North-West
NWGr.	North-West Greek
O	Old . . .
OAv.	Old Avestan (= Gatha-Avestan)
OBret.	Old Breton
OCS	Old Church Slavic
OCz.	Old Czech
OEng.	Old English
OHG	Old High German
OHitt.	Old Hittite
OIA	Old Indo-Aryan/Indic/Indian
OIc.	Old Icelandic
OInd.	Old Indic/Indo-Aryan/Indian (Vedic & Sanskrit)
OIr.	Old Irish
OLat.	Old Latin
OLatv.	Old Latvian
OLith.	Old Lithuanian
ONor.	Old Norse
OPers.	Old Persian
OPhyrg.	Old Phrygian
OPruss.	Old Prussian
ORuss.	Old Russian
OS	Old Saxon
Osc.	Oscan
OW	Old Welsh
OWCB	Old Welsh, Old Cornish and Old Breton
P	Proto . . .
Pal.	Palaic
Pamph.	Pamphylian
PAnat.	Proto-Anatolian
PBalt.	Proto-Baltic
PBSl.	Proto-Balto-Slavic
PCelt.	Proto-Celtic
Pers.	Persian
PGmc	Proto-Germanic
Phryg.	Phrygian
PIE	Proto-Indo-European
PIIr.	Proto-Indo-Iranian
Plb.	Polabian
Pol.	Polish
Pr.	Primitive
PSl.	Proto-Slavic
PToch.	Proto-Tocharian
Rom.	Romance
Russ.	Russian
RussCS	Russian Church Slavic
RV	Ṛg-Veda
S	South . . .

SCr.	Serbo-Croatian (= Bosnian/Croatian/Montenegrin/Serbian)
SE	South-East
Skr.	Sanskrit
Slav.	Slavic
Slk.	Slovak
Sln.	Slovene
Sogd.	Sogdian
Sp.	Spanish
Srb.	Serbian
Sw.	Swedish
Štok.	Štokavian
Thess.	Thessalian
Toch.	Tocharian (A & B)
Ukr.	Ukrainian
ULSorb.	Upper/Lower Sorbian
Umbr.	Umbrian
USorb.	Upper Sorbian
Ved.	Vedic
Ven.	Venetic
W	Welsh
W	West . . .
WGr.	West Greek
WSlav.	Western Slavic
YAv.	Young Avestan (= Late Avestan)

### **OTHER ABBREVIATIONS**

1 sg./pl.	first person singular/plural
2 sg./pl.	second person singular/plural
3 sg./pl.	third person singular/plural
abl.	ablative
acc.	accusative
act.	active
adj.	adjective
adm.	admirative
adv.	adverb
all.	allative
aor.	aorist
a.p.	accentual paradigm
arch.	archaic
caus.	causative
coll.	collective
comm.	common (gender)
comp.	comparative
cond.	conditional
dat.	dative
def.	definite
det.	determinate

---

dial.	dialect(al)
dim.	diminutive
du.	dual
encl.	enclitic
f.	feminine
fut.	future
gen.	genitive
imp.	imperative
imperf.	imperfective
impf.	imperfect
ind.	indicative
indef.	indefinite
indet.	indeterminate
inf.	infinitive
inj.	injunctive
instr.	instrumental
interrog.	interrogative
intr.	intransitive
IPA	International Phonetic Alphabet
loc.	locative
m.	masculine
med.	medial (= middle)
med.-pass.	medio-passive
mid.	middle (= medial)
n.	neuter
nom.	nominative
opt.	optative
pass.	passive
perf.	perfect
perfect.	perfective
pers.	person(al)
pl.	plural
pluperf.	pluperfect
pres.	present
pret.	preterite
pron.	pronoun
prv.	preverb
ptcl.	particle
ptcp.	participle
sg.	singular
sigm.	sigmatic
st.	stem
subj.	subjunctive
suff.	suffix
sup.	supine
them.	thematic
trans.	transitive
voc.	vocative



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

# INDO-EUROPEAN LANGUAGES – INTRODUCTION

*Mate Kapović*

Indo-European languages were originally, a couple of millennia ago, spoken from Europe to India. Now, in postcolonial times, they are spoken all over the world. English is thus, besides being a global *lingua franca*, spoken in North America, Australia, parts of Africa, and elsewhere; Spanish in Central and South America (Portuguese in Brazil); French in parts of Africa, etc. According to *Ethnologue*,<sup>1</sup> 437 IE languages (the number is in reality approximate) are spoken by almost 3 billion people in the world. With regard to the number of speakers, this makes the IE language family the biggest in the world. Concerning the number of individual languages, IE is not the largest language family – for instance, the African Niger-Congo family with more than 1500 languages is much larger. In the world, there are around 7000 languages (this is not an exact number, since it is extremely difficult to define linguistically what a language actually is), so in this sense IE languages are less important, although IE is still one of the biggest language families. Historical reasons – primarily the concentration of economic power in Europe and later in North America, together with the culture, science, etc., that goes with it – make IE the best-known language family in the world.

In the territories from Europe to India, where IE languages were spoken already a few thousand years ago, not all languages are IE. Most of the languages spoken in Europe today are IE, the exceptions (disregarding Balkan Turkish, the languages of the Caucasus, and recent migrant languages) being Basque, Maltese (an offspring of Arabic), and the Uralic languages (Hungarian, Estonian, Finnish, and a few smaller ones). Basque is the only remnant of pre-IE European languages. Other such cases (like Etruscan) have long disappeared. In the Middle East, we find both IE (like Persian/Farsi, Pashto, Kurdish, and Tajik) and non-IE languages (like Turkish, Azeri, Arabic, etc.). In the Indian subcontinent, many languages are IE (like Hindi/Urdu, Bengali, Punjabi, and Nepali), while many belong to the non-IE Dravidian family (like Tamil and Malayalam) and other non-IE language families (like smaller languages from the Tibeto-Burman and Austroasiatic families).

What is a language family? What does it mean when we say that English and Albanian belong to the IE language family? To put it succinctly, a language family consists of all the languages that have evolved from a single original proto-language. It is thus a group of genetically related languages (up to a certain point – it just may well be that all human languages are related, though this cannot be proven). In the IE case, it means that we can prove that all the IE languages that we know have evolved through many stages from a single parent language that we today call Proto-Indo-European. This proto-language was



**MAP 1.1** IE LANGUAGES TODAY

Source: Adapted from C. Watkins *The Indo-European Languages* (Routledge) 1st ed (1998)

spoken some 6000 years ago and, like most proto-languages, is unattested (which means that we have no written records of PIE). This is not always the case; for instance, Latin (as the proto-language of all Romance languages) and Old Chinese (as the proto-language of all modern Chinese “dialects”/languages) are attested.

How can one language split into more than 400 languages? The answer is simple – through language change. All languages change all the time (even today with all our schools, means of communication, and mass media). Linguistic changes that occur during the life of an individual speaker are usually slow and gradual (although still perceivable, whether they have to do with changes in pronunciation, grammatical forms, or vocabulary), but in time languages change enough to become mutually incomprehensible and sometimes very different. This is made easy by migration – when two social groups separate, it is easy to imagine that a previously uniform language can evolve in different directions (cf. the differences between British, American, and Australian English, although they are still in contact, mutually comprehensible, and thus considered one language).

That languages like English, Czech, and Welsh have all developed from a single proto-language was not always apparent. Already in ancient times, people noticed that there are many similar words in, for instance, Latin and Greek – cf. e.g. Lat. *māter* ‘mother’, *pater* ‘father’ and Gr. μήτηρ /mētēr/, πατήρ /patēr/ (or, for that matter, English *mother* and *father*). However, they did not know how to properly explain such similarities. Thus, there have been many unscientific hypotheses, such as that all languages are derived from Hebrew (considered a holy language) and so on. In the 15th and 16th centuries, there were a number of scholars, like Rodolphus Agricola/Roelof Huisman (1443/4–1485), Sigismund Gelenius/Zikmund Hrubý z Jelení (1497–1554), and Johann Elichmann (1601/2–1639), who noted correspondences between various IE languages. However, it is Marcus van Boxhorn (1612–1653), a professor at Leiden University, who

can be regarded as the first historical linguist and the father of IE linguistics. In 1647 he advanced a theory that Greek, Latin, Persian, Old Saxon, Dutch, German, Gothic, Russian, Danish, Swedish, Lithuanian, Czech, Croatian, and Welsh (but not Hebrew) originally stemmed from “Scythian” (= PIE). His friend Claudius Salmasius/Claude de Saumaise (1588–1653) added Sanskrit to the list. Amazingly, van Boxhorn also noted the importance of recognizing false cognates, loanwords, systematic correspondences, plausible semantic agreement, morphological comparison, and synchronically irregular forms (all basic elements of historical linguistics now called the “comparative method”).<sup>2</sup> The European discovery of Sanskrit played an important role in the beginnings of IE historical and comparative linguistics. Already in 1585, Filippo Sassetti (1540–1588), a merchant and scholar, noted in a private letter the similarity of some words in Sanskrit and Italian (e.g., Skr. *nava* and Ital. *nove* ‘nine’), and in the 18th century a couple of scholars wrote on the obvious similarity of Sanskrit to European languages. One of them was William Jones (1746–1794), who in a lecture in 1786 compared Sanskrit to Greek, Latin, Gothic, Celtic, and Old Persian, claiming that they “have sprung from some common source.” Jones is usually wrongly considered the founding father of IE historical linguistics, although he was neither the first to make such a hypothesis nor the one with the clearest presentation of the problem. Shortly after, the first modern IE historical linguists – Jacob Grimm (1785–1863; famed as one of the Grimm brothers), Rasmus Rask (1787–1832), and Franz Bopp (1791–1867) – started doing serious historical linguistics, and others followed. The very term *Indo-European* was coined in 1813 by Thomas Young (1773–1829). The discovery of Anatolian and Tocharian texts in the beginning of the 20th century gave new impetus for IE linguistics.

The IE language family comprises ten principal branches (those being the ones sufficiently attested in ancient times and/or today) and a number of fragmentarily attested separate languages spoken in ancient times (like Lusitanian or Phrygian) that have since disappeared. Three main branches (Anatolian, Indo-Iranian, and Greek) are attested already in the second millennium BCE, two in the first millennium BCE (Italic, Celtic), and the other five branches later. Three of the principal branches consist of one language only (Greek, Armenian, and Albanian), although those comprise different dialects as well. Two major branches (Anatolian and Tocharian) are now extinct, and the same goes for all fragmentarily attested languages and certain languages from other primary branches (like Gaulish). The ten principal IE branches are as follows.

1. **Anatolian** – attested from the 19th/16th to the 1st century BCE

Spoken, until its disappearance, in Asia Minor. Old Anatolian languages (spoken in the 2nd millennium BCE) are Hittite (the best attested, written in Akkadian cuneiform, the language of the great Hittite Empire), Cuneiform Luwian (Luvian), Hieroglyphic Luwian, and Palaic. Hittite is attested from around the 16th century, but a few names and loanwords are found already in the 19th century BCE in Old Assyrian texts. New Anatolian languages (spoken in the 1st millennium BCE) are Lydian, Lycian (Lycian A), Milyan (Lycian B), Carian, Sidetic, and Pisidian.

2. **Indo-Iranian** – attested from the 15th–14th century BCE

Spoken mostly in the northern part of the Indian subcontinent and the Middle East (Eastern Turkey, Caucasus, Iran, Afghanistan). As a result of old migrations, the Romani language dialects are spoken by the Roma people in Europe. Indo-Iranian consists of two great groups – Indo-Aryan (Indic) and Iranian. The third, smaller group is Nuristani (on the border between Afghanistan and Pakistan), and the fourth, disputably, is Dardic (traditionally considered part of Indo-Aryan). The two major



groups are attested from ancient times and represented by many languages, while the two small groups are attested only from modern times and consist of a few small languages spoken in remote parts of northern India, Afghanistan, and Pakistan. Indo-Aryan languages are chronologically divided into Old Indic/Indo-Aryan (Vedic – the language of the Vedas; Sanskrit – the classical language of India), Middle Indic/Prakrits (Pali, Shauraseni, Maharashtri, Magadhi, etc.), and New Indic (many languages – Hindi/Urdu, Bengali, Punjabi, Marathi, etc.). Certain words and names of Indo-Aryan origin were attested in the 15th–14th century BCE in Hurrian texts. The Rigveda (the oldest part of the Vedas) is usually dated to the second half of the 2nd millennium BCE but was transmitted orally for a long time. Iranian languages are chronologically divided into Old Iranian (Avestan – the language of Zarathustra, from the first half of the 1st millennium BCE, or perhaps even earlier but at first transmitted orally; Old Persian – the language of the Achaemenid Empire), Middle Iranian (Pahlavi – the language of the Sassanid Empire; Sogdian; etc.), and New Iranian (New Persian/Farsi – the official language of Iran; the Kurdish languages; Pashto – the official language of Afghanistan; etc.).

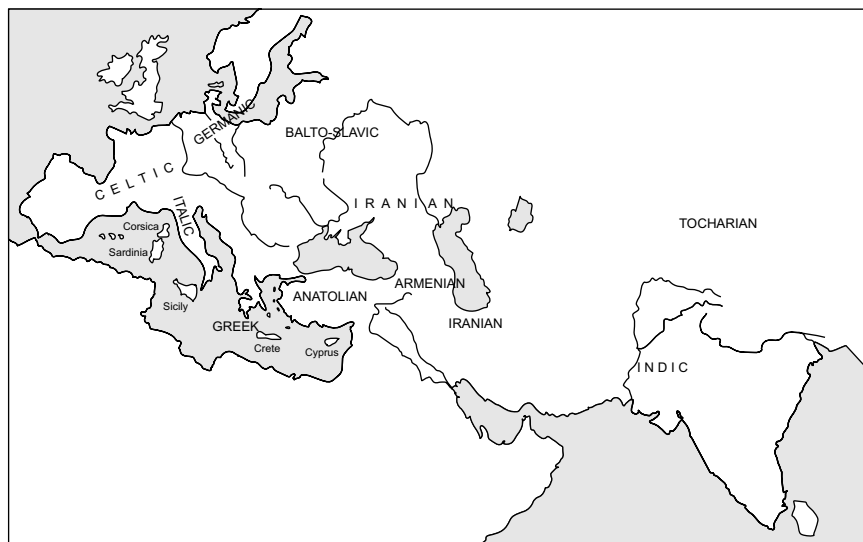
3. **Greek** – attested from the 15th–14th century BCE  
Spoken from ancient times until the present in mainland Greece and neighboring islands. The first attestations were written in the Linear B script in Mycenaean Greek, the earliest recorded Greek dialect. From the 8th century, Greek has been written in the Greek alphabet. In the classical period (around the 5th century BCE), the following Greek dialects existed – North-West dialects, Doric, Aeolic (controversial), Arcado-Cypriot, and Ionic-Attic (the details of the subgrouping are disputed). Later stages of Greek are Hellenistic/Koiné, Byzantine/Medieval, and Modern Greek.
4. **Italic** – attested from the 7th century BCE  
Spoken in ancient Italy. Italic consists of two groups – Latino-Faliscan and Sabellian (Osco-Umbrian). The Latino-Faliscan group consists of Latin (very well attested) and Faliscan. The Sabellian group consists of Oscan, Umbrian, and a number of smaller and poorly attested Italic languages. All Italic languages except Latin disappeared already in Roman times owing to the expansion of Latin as the language of the Roman state. Latin later developed into numerous modern Romance languages (Portuguese, Spanish, Catalan, French, Italian, Romanian, etc.).
5. **Celtic** – attested from the 6th century BCE  
Now spoken in peripheral areas of the British Isles (Ireland, Wales, Scotland) and France (Brittany), previously spoken in wide areas of Europe (from the Iberian Peninsula and France all the way to Asia Minor). The Celtic branch can be divided (though controversially) into three groups: Celtiberian, Continental Celtic, and Insular Celtic. Celtiberian is a one-language group, attested in the 2nd to 1st century BCE in present-day Spain. Continental Celtic languages – Gaulish (in France) and Lepontic (in northern Italy) – are now extinct. Insular Celtic consists of Brythonic (Welsh, Breton, and extinct Cornish) and Goidelic (Irish/Gaelic, Scottish Gaelic, and extinct Manx). Of all currently existing Celtic languages, only Welsh has a more or less secure future.
6. **Germanic** – attested from the 3rd (?) century BCE  
Nowadays spoken in North-West Europe. The first Germanic words are attested perhaps already in the 3rd (?) century BCE on the so-called Negau helmet (found in

Ženjak in present-day Slovenia). The earliest runic inscriptions date from the first centuries CE, and the Gothic translation of the Bible from the 4th century. Germanic languages comprise three groups: East, North, and West. The extinct East Germanic group consists of Gothic, Vandal, and Burgund. The Northern languages are Old Norse and, in modern times, the Scandinavian languages. The Western languages are Old High German, Old Saxonian, Old Low Franconian (Old Dutch), Old Frisian, and Old English (in modern times, respectively, German, Low German, Dutch, Frisian, and English).

7. **Armenian** – attested from the 5th century CE  
Spoken in present-day Armenia and elsewhere in the Caucasus and Middle East as a minority language, previously also in wide areas of present-day eastern Turkey. Modern Armenian consist of two dialects, Eastern and Western, both standardized.
8. **Tocharian** – attested from the 4th–5th to the 8th–10th century CE  
Spoken in merchant cities on the ancient Silk Road, in the present-day Xinjiang province of western China (now inhabited by Turkic Uyghurs). There were two closely related Tocharian languages – Tocharian A and Tocharian B.
9. **Balto-Slavic** – attested from the 9th (Slavic) and 16th (Baltic) century CE  
Spoken in the Baltic, the Balkans, and Central and Eastern Europe. The Balto-Slavic branch consists of West Baltic, East Baltic, and Slavic. The extinct Old Prussian is West Baltic, and Lithuanian and Latvian are East Baltic (there are a few other poorly attested and extinct Baltic languages). Slavic consists of East (Russian, Belorussian, Ukrainian), West (Polish, Czech, Slovak, etc.), and South Slavic (Slovene; Bosnian/Croatian/Montenegrin/Serbian – traditionally known as Serbo-Croatian; Macedonian; Bulgarian). The first attested Slavic language, Old Church Slavic, was based on the Thessaloniki Old Macedonian dialect.
10. **Albanian** – attested from the 15th century CE  
Spoken in present-day Albania, Kosovo, West Macedonia, South Montenegro, North-West Greece, and South Italy. There are two dialects – Gheg in the north and Tosk in the south, the latter being the basis of Standard Albanian.

Fragmentarily attested IE languages can be divided into two basic groups: those that have at least one attested text/inscription and those that are attested only through onomastics or individual words in texts written in other languages. In the first group, there are Lusitanian (Portugal), Venetic (North-East Italy), Messapic (the “heel” of Italy), Thracian (Bulgaria), and Phrygian (Minor Asia). In the second group, we find languages like Ligurian (North-West Italy), the Illyrian languages (North-West Balkans), Dacian (Romania), and Macedonian (no relation to present-day Slavic Macedonian). Apart from fragmentarily attested IE languages, there are also possible non-attested IE languages, whose traces can perhaps be seen in loanwords in other languages (like the supposed North-West Block language in Benelux), but these are rather speculative.

Many elements can be reconstructed for the last stage of PIE. However, what we obtain through reconstruction will never be a full picture of the language, since elements (sounds, grammatical endings, words, meanings, etc.) can disappear with no trace or leave so few traces that it is not possible to reconstruct them plausibly. Thus, the around 1500 PIE lexical items that have been reconstructed are certainly just a small part of the actual PIE vocabulary, which must have had something like 30,000–50,000 words, just



**MAP 1.2** IE LANGUAGES IN THE 1ST MILLENNIUM BCE

Source: Adapted from C. Watkins *The Indo-European Languages* (Routledge) 1st ed (1998)

as much as any language today.<sup>3</sup> Likewise, we can, for instance, reconstruct the ending variants \*-mes, \*-mos, \*-mesi, \*-men, etc. in the PIE 1st person plural present indicative (p. 93), but it is impossible to know how to interpret them exactly. Were these just random variants (not likely), were they sociolinguistic variables, did their use depend on dialect, gender, social status, etc.? Sometimes it is also very difficult to separate different layers – should an element be reconstructed for PIE in general, dialectal PIE, post-Anatolian PIE (i.e., PIE after the early split of Anatolian), early PIE (i.e., PIE including Anatolian), post-PIE time, etc.?

Like most languages that cover any substantial ground, PIE probably also had dialects. The dialectal picture of PIE is not completely clear because it is sometimes difficult to distinguish PIE dialectal features from post-PIE shared innovations and accidental independent innovations. It is now widely accepted that Anatolian was the first branch to separate from the rest of PIE (and that Tocharian was the next). However, which Anatolian characteristics (like the lack of feminine gender) are archaisms, and which are innovations, is still a matter of dispute (p. 62, 66, 159, 178). Dialects can share one trait with one dialect, and another with another dialect. For instance, Greek, Indo-Iranian, and Armenian all have the augment \*h<sub>1</sub>e- in past tense verbal forms, which is not present in other IE languages. At the same time, Armenian and Indo-Iranian are satem languages, while Greek is a centum language (p. 27). In Greek, Armenian, and Phrygian, initial pre-consonantal laryngeals (\*H-) yield vowels (p. 48), while this is not the case in Indo-Iranian. Indo-Iranian, on the other hand, shares the so-called RUKI rule (p. 206) with Balto-Slavic (p. 29, 495, 526–527), a satem branch, while Balto-Slavic shares the ending \*-mos in the dative plural with Germanic (as opposed to \*-b<sup>h</sup>os in Italic, p. 63), a centum branch, etc.

The question of the PIE homeland (the “original” territory where speakers of PIE lived) is still a subject of debate. The most commonly accepted hypothesis in the last few decades has been the “Kurgan hypothesis” (Kurgans being a type of burial mound), which

puts Proto-Indo-Europeans in the steppes of present-day Ukraine and South Russia, north of the Black Sea. The position of the PIE homeland is often deduced from certain linguistic indications, and by trying to connect linguistic evidence with archeology, paleobotany, and other disciplines. The way linguistic evidence is used in trying to deduce the place of the PIE homeland follows a certain logic. For instance, since one can reconstruct a PIE word for ‘snow’ (\**snoyg*<sup>whs</sup>), we can suppose that PIE was spoken in a place where there was snow (which by itself does not amount to much). Evidence has to be carefully evaluated – for instance, Lat. *mare*, OCS *morje*, and OIr. *muir* ‘sea’ would all point to PIE \**mori* ‘sea’ (and thus to a PIE homeland by the sea). However, Goth. *marei* (cf. Eng. *mere*, *marsh*) means both ‘sea’ and ‘lake’, which brings the PIE meaning ‘sea’ in doubt (it is not unimaginable that the original meaning was ‘lake, standing water’ and that the meaning ‘sea’ is secondary, though there can be no certainty in any of the possibilities).

Comparative linguistics has proven beyond reasonable doubt that all IE languages stem from the same proto-language. However, the question of a deeper genetic relationship of IE with other language families is still a matter of heated debate (though most IE-ists, curiously, ignore the problem, choosing not to go outside of inner-IE problems). The most likely IE relative is Uralic. There are some striking cognate-looking examples, cf. e.g. Proto-Uralic \**weti* ‘water’ and PIE \**wodǵ* (this is not a word easily loaned), or Proto-Uralic verbal endings 1 sg. \**-mi*, 1 pl. \**-me*, 2 pl. \**-te*, which are practically identical to PIE \**-m(i)*, \**-me*, \**-te* (p. 93). Although the Indo-Uralic hypothesis is still not strictly proven, grammatical correspondences like these make this proposition quite probable. On a wider scale, IE often connected with “Nostratic” languages, which include IE, Uralic, Kartvelian (like Georgian), Altaic (a suspicious family by itself, consisting of Turkic, Mongolic, Tungusic, and possibly other languages), Dravidian, Afro-Asiatic (consisting of Semitic, Egyptian, etc.), and possibly other families/languages (the “cast” is not always the same). There is a lot of work left to be done in long range comparison, but it is not at all certain that we will ever get satisfactory answers to these problems. Many linguists are, with good reason, highly skeptical concerning the very methodological basis of such reconstruction, and simply do not believe that there is enough linguistic material preserved (since sounds, forms and words not only change but can also disappear leaving no trace) to let us plausibly reconstruct proto-languages so distantly in the past.

Up till now, we have seen examples of cognate words and reconstructed PIE forms (marked with an asterisk to indicate them not being attested). But how do linguists actually know what words/forms to compare, and how do they get the idea that two or more languages are genetically related in the first place? Looking for similar words with similar meanings is obviously not enough, since similar words with similar meanings in two languages can easily be the result of linguistic borrowing. Compare Modern English and French pairs like *city/cité*, *people/people*, or *part/part*. These are all French loanwords in English, which can be proven by historical data (French loanwords were brought to England by French-speaking Normans after their 1066 conquest). When looking for genetically related languages, it is safer to look for similarities in grammar (morphology, i.e., nominal or verbal endings, etc.), since morphological elements are borrowed much less frequently than vocabulary. Looking for morphological similarities may be a problem in languages which are not rich with morphology (like Chinese or even English), but not in older IE languages, which have quite complex morphology. Some of the similarities, clearly seen in ancient IE languages, are visible even today. For instance, compare the plural ending *-i* in Italian *lupi* ‘wolves’ and Russian *волки* /*vólki*/. However, one does need to reconstruct vocabulary as well. When looking for old words inherited directly from PIE, one has to find words that are likely to have been in a language already a

couple of millennia ago (thus, not words like *computer* or *tennis*) and words that are not so likely to be borrowed. Although in principle any word can be borrowed (and all kinds of examples from all around the world can be found), in some types of vocabulary one often finds old inherited words, for instance, in words for family relations (cf. Eng. *father*, *mother*, *brother* and related Ital. *padre*, *madre*, *fratello*), numbers (cf. Eng. *two*, *three*, *seven* and related Ital. *due*, *tre*, *sette*), body parts (Eng. *nose*, *beard*, *ear* and Ital. *naso*, *barba*, *orecchio*), etc. Genetically related words are sometimes obviously similar (Eng. *mother* ~ Ital. *madre*), sometimes less obviously so (Eng. *wolf* ~ Ital. *lupo* – the *l* is the same, and the sounds *f* and *p* are phonetically close), and sometimes almost completely different (Eng. *hornet* ~ Ital. *calabrone*). Of course, similarity of form and meaning may be completely accidental, cf. e.g. non-related pairs such as Eng. *much* and Span. *mucho*, Eng. *iron* and Span. *hierro*, Eng. *day* and Span. *día*, etc. (this is more apparent when one considers phonologically more archaic Germanic cognates such as Scots *mickle*, German *Eisen*, or Swedish *dag* respectively, or the Latin forms *multum*, *ferrum*, *diēs* that yield the cited Spanish forms).

But similarity alone is not enough. Not only does it not guarantee that two words are related, it is also not scientific. What historical linguists look for are not necessarily similar sounds, words, or forms, but systematic sound correspondences. This means that certain phonemes (or clusters of them) in purportedly genetically related languages have to systematically correspond to each other, whether they are phonetically close or not. For instance, in initial position we find in English and Italian the correspondences *n-* ~ *n-* (*no/no*, *night/notte*, *nine/nove*), *h-* ~ *c-* (*hornet/calabrone*, *hound/cane*, *heart/cuore*), *t-* ~ *d-* (*two/due*, *ten/dieci*, *tooth/dente*), etc. Careful examination and analysis of such correspondences is one of the primary tasks of historical linguistics and a prerequisite for proving a genetic relationship between two or more languages. Describing such correspondences will be an important part of the next chapter.

**Note:** I would like to thank David Mandić and Petra Šoštarić for reading the first draft of the chapter. Of course, all the mistakes are just mine.

## FURTHER READING

Lehmann 1967 provides a rich selection of translated primary texts (with commentary) from the beginnings and early IE historical linguistics. Pedersen 1959 is also still very useful. For the problems of the IE homeland, culture, and archeology from a mainstream IE-ist perspective see, for instance, Mallory 1989 and Mallory & Adams 1997. A lot of information on IE culture can be found in the famous but controversial and idiosyncratic Gamkrelidze & Ivanov 1995. For more on the “Kurgan hypothesis” from the pen of its originator see Gimbutas 1997. Salmons & Joseph 1998 is an informative look at Nostratic from different perspectives. There are a number of good handbooks and introductions to historical linguistics, many of them using a lot of IE material. Trask 2007 may be a good place to start, as well as Campbell 2004. Joseph & Janda 2003 and Ringe & Eska 2013 are a bit more demanding. Labov’s trilogy on the principles of linguistic change (1994, 2001, 2010), written mostly (but not exclusively) from the perspective of quantitative sociolinguistics, is invaluable and inspiring, though perhaps not a usual IE-ist read. For a list of recent introductions to IE, see the “Further reading” sections in the chapters on PIE phonology and PIE morphology in this volume.

## NOTES

- 1 <http://www.ethnologue.com/statistics/family> [last approached: Feb. 2, 2016].
- 2 For van Boxhorn and other early scholars cf. Van Driem 2001: 1039–1051.
- 3 Cf. Mallory & Adams 2006: 117–119.

## REFERENCES

- Campbell, Lyle 2004, *Historical Linguistics. An Introduction*, Edinburgh University Press, Edinburgh/MIT Press, Cambridge
- Gamkrelidze, Thomas V. & Ivanov, Vjačeslav V. 1995, *Indo-European and the Indo-Europeans. A Reconstruction and Historical Analysis of a Proto-Language and a Proto-Culture*, Mouton de Gruyter, Berlin–New York
- Gimbutas, Marija 1997, *The Kurgan Culture and the Indo-Europeanization of Europe. Selected Articles from 1952 to 1993*, Robbins Dexter, Miriam & Jones-Bley, Karlene (eds.), Institute for the Study of Man, Washington, DC
- Joseph, Brian D. & Janda, Richard D. (eds.) 2003, *The Handbook of Historical Linguistics*, Blackwell Publishing, Oxford
- Labov, William 1994, *Principles of Linguistic Change. Volume 1: Internal Factors*, Blackwell, Oxford–Cambridge
- Labov, William 2001, *Principles of Linguistic Change, Volume II: Social Factors*, Blackwell, Oxford–Cambridge
- Labov, William 2010, *Principles of Linguistic Change. Volume 3: Cognitive and Cultural Factors*, Wiley-Blackwell, Oxford
- Lehmann, Winfred P. 1967, *A Reader in Nineteenth Century Historical Indo-European Linguistics*, Indiana University Press, Bloomington–London
- Mallory, James P. 1989, *In Search of the Indo-Europeans: Language, Archeology and Myth*, Thames & Hudson, London
- Mallory, James P. & Adams, Douglas Q. (eds.) 1997, *Encyclopedia of Indo-European Culture*, Fitzroy Dearborn Publishers, London–Chicago
- Mallory, James P. & Adams, Douglas Q. 2006, *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*, Oxford University Press, Oxford–New York
- Pedersen, Holger 1959, *The Discovery of Language: Linguistic Science in the 19th Century*, Indiana University Press, Bloomington
- Ringe, Don & Eska, Joseph F. 2013, *Historical Linguistics. Toward a Twenty-First Century Reintegration*, Cambridge University Press, Cambridge
- Salmons, Joseph C. & Joseph, Brian D. (eds.) 1998, *Nostratic: Sifting the Evidence*, John Benjamins, Amsterdam–Philadelphia
- Trask, Robert L. 2007<sup>2</sup>, *Trask's Historical Linguistics*, revised by Robert McColl Millar, Hodder Arnold, London
- Van Driem, George 2001, *Languages of the Himalayas: An Ethnolinguistic Handbook*, vol. II, Brill, Leiden



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

PART 1

# **PROTO-INDO-EUROPEAN**





Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

# PROTO-INDO-EUROPEAN PHONOLOGY

*Mate Kapović*

## INTRODUCTION

Proto-Indo-European phonology is reconstructed on the basis of its reflexes in Indo-European languages – mostly older ones, generally preserving more archaic features, i.e., closer to PIE. The principal languages referenced in this chapter for the reconstruction of PIE are Old Indic/Indo-Aryan (Vedic and Sanskrit), Avestan (Old/Gatha- and Young/Late Avestan), (Ancient) Greek, Latin, Old Church Slavic, Lithuanian (Old and Modern), Gothic, Hittite, Old Irish, Tocharian (A and B), (Classical) Armenian, and Albanian. Of course, other languages (Sabellic, Latvian, Old Prussian, Old High German, Old English, Luwian, Gaulish, Old Persian, etc.) are also used when necessary or when they have a form or feature not present in a language from the first group. Ideally, PIE reconstruction should take into account all the relevant information present in any IE language/dialect. This overview tries to present the reconstructed PIE phonological system and its main reflexes in major ancient IE languages. A more detailed overview of reflexes for specific languages/branches can be found in separate chapters in this volume.

The following phonological system can be reconstructed for the last phase of PIE:

STOPS	*p	*t	*k	*k̥	*kʷ	[voiceless]
	(*b)	*d	*g	*g̊	*gʷ	[voiced]
	*bʰ	*dʰ	*gʰ	*g̊ʰ	*gʷʰ	[aspirated]
FRICATIVES	*s	*h₁	*h₂	*h₃		
RESONANTS and GLIDES	*m/ṃ	*n/ṅ	*l/ḷ	*r/ṛ	*w/u	*y/i
VOWELS (including DIPHTHONGS)	(*a)	*e	*o			
		*ē	*ō			
	(*ay)	*ey	*oy	(*aw)	*ew	*ow
		*ēy	*ōy		*ēw	*ōw

The phonemes \*b and \*a (together with \*ay and \*aw) can be considered marginal, or at least infrequent, phonemes (p. 18, 42), \*a also being an allophone of \*e when in contact with \*h₂ (p. 43). Reconstruction of \*ā not stemming from original \*eh₂ is very questionable (p. 42). Short \*i and \*u most likely had long allophones in monosyllabic forms (p. 55).

PIE stops were voiceless, voiced, or (voiced) aspirated. The velar stops were unmarked (\*k, \*g, \*g<sup>h</sup>), palatalized (\*k̟, \*ġ, \*ġ<sup>h</sup>) (IPA [k<sup>j</sup>, g<sup>j</sup>, g<sup>jh</sup>]), or labialized (\*k<sup>w</sup>, \*g<sup>w</sup>, \*g<sup>wh</sup>). Besides \*s, the “laryngeals” \*h<sub>2</sub> and \*h<sub>3</sub> (marked numerically) were almost certainly fricatives as well (some kind of *h*-type sounds). The \*h<sub>1</sub> “laryngeal” (provisionally labeled a fricative in the above scheme) might also have been a fricative, but this is less certain (p. 30). The sign \*H is used for all three laryngeals when it is not known which one is to be reconstructed or when we need a mark for all three laryngeals. Resonants and semivowels were either asyllabic (\*m, \*n, \*l, \*r, \*w, \*y (IPA [j])) or syllabic (\*ṁ, \*ṅ, \*ṁ̥, \*ṅ̥, \*u, \*i) (IPA [ṁ, ṅ, l̥, r̥]). The vowels \*i and \*u act phonologically in the same way as syllabic resonants – the relation of \*i/u to \*y/w is the same as that of \*ṁ/ṅ/l̥/r̥ to \*m/n/l/r. Some reconstructable segments appeared only as allophones – cf. the non-phonemic fricatives \*z (p. 29), \*þ (IPA [θ]), and \*ð (and \*ð<sup>h</sup>) (p. 53) and the vowel \*ə (p. 46–47, 52).

What is immediately typologically noticeable in the PIE phonological system, among other things, is the paucity of “real” vowels (if one does not take into account syllabic glides, long vowels, and diphthongs), the very unusual marginality (or even complete lack of) the vowel *a*, and the complete lack of affricates (no phonemic \*c [tʃ], \*č [tʃ], \*ʒ [dʒ], etc.). On the other hand, the stop system was rather rich (with 15 of them). In later IE languages, phonological systems usually developed in the following direction: the vowel system became typologically more common, with *a* becoming more frequent (additionally, *i* and *u* lost their phonological relation to *j* and *w/v*, while diphthongs often became monophthongs); syllabic resonants (\*ṁ, \*ṅ, \*ṁ̥, \*ṅ̥) disappeared everywhere (later they appeared again in some languages); a lot of languages developed new affricates; the “laryngeals” (\*h<sub>1</sub>, \*h<sub>2</sub>, \*h<sub>3</sub>) disappeared in most languages in most positions, but new fricatives appeared; the stop system simplified everywhere; and languages preserved only one additional distinctive feature – aspiration or labialization. Velar palatalization as such was nowhere preserved – it either completely disappeared or was the origin for new affricates and fricatives.

Note on the PIE “orthography”: various authors/schools write reconstructed PIE phonemes differently. Palatalized velars can also be written as \*kʲ, \*k̟, \*k<sup>j</sup>, \*k<sup>y</sup>, etc.; \*k<sup>ɥ</sup> can be written instead of \*k<sup>w</sup>; aspirated stops can be written as \*bh, \*dh, \*gh; laryngeals can be written as \*H<sub>1</sub>, \*H<sub>2</sub>, \*H<sub>3</sub> or \*ə<sub>1</sub>, \*ə<sub>2</sub>, \*ə<sub>3</sub> (the latter usually for “syllabic” laryngeals), and an unknown laryngeal as \*h<sub>x</sub> (instead of \*H); the glides \*y, \*w can be written as \*i/j, \*u; \*eu/eu instead of \*ew, etc.

TABLE 1.1 PIE STOPS

	voiceless	voiced	aspirated
labials	*p	*b	*b <sup>h</sup>
dentals	*t	*d	*d <sup>h</sup>
velars	*k	*g	*g <sup>h</sup>
palatalized velars	*k̟	*ġ	*ġ <sup>h</sup>
labialized velars	*k <sup>w</sup>	*g <sup>w</sup>	*g <sup>wh</sup>

No modern IE language has preserved PIE palatalized velars as such. On the other hand, labiovelars are preserved even in some modern IE languages – cf. the phonetics of English *queen* [k<sup>w</sup>i:n] (p. 25). PIE aspirated stops are traditionally reconstructed as

voiced because of their reflexes, which are mostly voiced (p. 28). However, voiced aspirated stops are typologically less common than voiceless aspirated stops – cf. the initial allophonic aspirated [p<sup>h</sup>], [t<sup>h</sup>], [k<sup>h</sup>] in English *pin*, *tin*, *kin*. In PIE, velars could be both palatalized and aspirated (\*g<sup>h</sup>) and labialized and aspirated (\*g<sup>wh</sup>) at the same time.

PIE voiceless stops do not change in most of the languages (*p*, *t*, *k*), the main exceptions being Germanic (where they change to voiceless fricatives *f*, *þ*, *x*) and Armenian (where they change to voiceless aspirates \*p<sup>c</sup>, t<sup>c</sup>, k<sup>c</sup>). The phoneme \*p weakens and disappears in Celtic. In most languages, PIE voiced stops remain unchanged (*b*, *d*, *g*). However, in Germanic and Armenian (perhaps also in Anatolian) they become voiceless *p*, *t*, *k* (in Tocharian all stops merge into *p*, *t*/*s*, *k*). The phonemes traditionally reconstructed as voiced aspirated stops change in most IE languages. They are preserved only in Indo-Aryan (*bh*, *dh*, *gh*); in Greek they remain aspirated but become voiceless (φ [p<sup>h</sup>], θ/θ [t<sup>h</sup>], χ [k<sup>h</sup>]); in Italic they change to fricatives (voiceless initially: \*f-, \*þ-, \*x-; voiced medially: \*-β-, \*-ð-, \*-γ-), and in other branches they change to plain voiced stops (*b*, *d*, *g*). PIE palatovelars change to plain velars in some languages (*k*, *g*) and to affricates (c[ts]/ć[tc]/č[ʧ], dʒ[dʒ]/dʒ̣[dʒ̣]/dʒ̣̣[dʒ̣̣] . . .) and, further on, to fricatives (s/š[ç]/ś[ʃ], z/ž[z]/ẓ̌[ʒ] . . .) in others. PIE labiovelars remain labialized velars in some languages (k<sup>w</sup>, x<sup>w</sup>, g<sup>w</sup>), or change further to full labials (*p*, *b*/*w*), while in others (mostly those where PIE palatovelars yield affricates/fricatives) they become plain velars (*k*, *g*).

## Voiceless stops

### \*p

**PIE \*pōds** ‘foot’ > Ved. *pāt* (YAv. *pad-*), Gr. ποῦς (cf. via Lat. Eng. *podium*), Lat. *pēs* (⇒ Eng. *pedal*), OCS *pěšb* ‘on foot’, Lith. *pėsčias* ‘pedestrian’, Goth. *fofus* (Eng. *foot*), OIr. *is* ‘below, under’ < \*pēdsu, Hitt. *pat-*, Arm. *otn* ‘foot’, *het* ‘foot(print)’, Toch. B *paiyye*

**PIE \*pek<sup>w</sup>-** ‘cook, bake’ > Ved. *pācanti* ‘they cook’ (YAv. *pačata* ‘it is cooked’), Gr. πέσσω, Lat. *coquō* < \*k<sup>w</sup>ek<sup>w</sup>- (assimilation), OCS *pešti*, Lith. *kèpti* < \*pekti (metathesis), MW *pobi* < \*k<sup>w</sup>ok<sup>w</sup>- (assimilation), Arm. *hac* ‘bread’, Alb. *pjek*, Toch. AB *pāk-*

**PIE \*peh<sub>2</sub>wr/n-** ‘fire’ (p. 78) > Gr. πῦρ, Umbr. *pir*, Cz. (arch.) *pyr* ‘ash, fire’, OPruss. *panno*, Goth. *fon* (Eng. *fire*), Hitt. *paḫhur*, Arm. *hur*, Toch. B *pūwar*

PIE \*p is reconstructed because it reflects as *p* in most languages. In Germanic, where all stops change their articulation, we find *f* (medially only following the stressed syllable). In Armenian the reflex is originally aspirated \*p<sup>h</sup>-, which yields either *h-* or nothing word-initially (#*h-* of any origin is unstable in Armenian) and *w* word-internally (p. 426–427). In Celtic PIE \*p disappears through a \*p > \*f > \*h type process (p. 363). In Italic and Celtic (before the \*p > \*h change in Celtic), \*p–k<sup>w</sup> (a \*p and a \*k<sup>w</sup> in the same word) assimilate regularly to \*k<sup>w</sup>–k<sup>w</sup>, cf. PIE \*penk<sup>w</sup>e ‘five’ > Lat. *quīnque*, OIr. *cóic* (\*k<sup>w</sup> > *c* in Old Irish). In Hittite the reflex is *p-* initially and a geminate spelling *-pp-* internally (p. 177).

### \*t

**PIE \*tod** ‘that’ > Ved. *tát* (GAv. *ta-*), Gr. τό, OCS *to*, Lith. *tàs*, Goth. *þata* (Eng. *that*, Germ. *das*), OIr. *tó* ‘yes’, Alb. *kë-ta*, Arm. *da*

**PIE \*t<sup>h</sup>h<sub>2</sub>us** ‘thin’ > Ved. *tanús*, Gr. *τανός*, Lat. *tenuis*, OCS *тънѣцаје* ‘thinner’, Lith. *tėvas*, OIc. *punnr* (Eng. *thin*, Germ. *dünn*), OIr. *tanai*

**PIE \*pet-** ‘fly’ > Ved. *pátati* ‘flies’ (YAv. *patənti* ‘they fly’), Gr. *πέτομαι* ‘I fly’, Lat. *petō* ‘I go after’, OEng. *feðer* < Gmc \*feþrō (Eng. *feather*), OIr. *én* ‘bird’ < \*pet-no- (Gaul. *Etnosus* (name of a deity)), Hitt. *pattai-*

PIE \*t is reconstructed because it reflects as *t* in most languages. In Germanic it changes to *b* (Eng. <th>, Gmc \*þ > *d*) initially and medially following the stressed syllable (p. 20). In Armenian the reflex is an aspirated *tʰ* (cf. Arm. *tʰaršamim* ‘I wither’, Lat. *torreō* ‘I dry’, Eng. *thirst* < PIE \*ters-), but the evidence is scarce, with *d-* appearing in unaccented words (cf. Arm. *du* ‘thou’, Lat. *tū* < PIE \*tū, p. 438). In Hittite we find initial *t-* and medial *-tt-*.

### \*k

**PIE \*lew<sub>1</sub>k-** ‘light’ > Ved. *rókás* (GAv. *raočaḥ-* with palatalization), Gr. *λευκός* ‘light, white’ (⇒ Eng. *leukocyte*), Lat. *lūx*, Russ. CSI. *lučb* (with palatalization), Lith. *laūkas* ‘pale’, Goth. *liuhap* (Eng. *light*), Mlr. *luach* ‘glowing white’ (Gaul. *Leucus* (pers. name)), Hitt. *lukk-* ‘to light up’, Toch. AB *luk-*

**PIE \*krew<sub>2</sub>h<sub>2</sub>s** ‘raw meat, blood’ > Ved. *kraviṣ-* ‘raw/bloody meat’, Gr. *κρέας* ‘meat’, Lat. *cruor* ‘(wound) blood’, OCS *krъnъ* ‘blood’, Lith. *kraūjas* ‘blood’, OEng. *hrēaw* (Eng. *raw*)

**PIE \*kes-** ‘comb, scratch’ > Ved. *kṣnaumi* ‘I whet’ (\*ksnew-), Gr. *ξέω* ‘I scrape’, PSI. \*kosa ‘hair’ (BCMS *kōsa*), Lith. *kasà* ‘braid’, OIr. *cír* ‘comb’ (also *cass* ‘curl’, Gaul. *Cassi-* (pers. name)), Hitt./CLuw. *kiš-* ‘to comb’, Arm. *kʰos* ‘scab, itch’, (?) Toch. B *kāswo* ‘eruption, inflammation of the skin’

PIE \*k is reconstructed because it reflects as *k* in most languages (cf. PIE \*kēs- > Alb. *kohë* ‘time’, OCS *časъ* ‘time, moment’). Parallel to \*p and \*t, \*k yields \*x in Germanic (medially only following the stressed syllable) and aspirated *kʰ* in Armenian. In Hittite we find initial *k-* and medial *-kk-* (written also as <gg>).

## Voiced stops

### \*b

**PIE \*h<sub>2</sub>ebōl** ‘apple’ > OCS *ablъko*, Lith. arch./dial. *óbuolas*, OEng. *æppel* (Eng. *apple*), OIr. *ubull*, (?) Lat. (from Osc.) *Abella* (name of a town)

**PIE \*bel-** ‘strong’ > Ved. *bálam* ‘strength’, Gr. *βέλτερος* ‘better’, Lat. *dē-bilis* ‘weak’, OCS *bolii* ‘better’, Low Germ. *pall* ‘unmovable’, (?) OIr. *balc* ‘strong’

**PIE \*d<sup>h</sup>ewb-** ‘deep’ > Gr. *βυθός* ‘depth’ (< \*θυβός, metathesis), OCS *dъbrъ* ‘valley, ravine’ < \*dъbrъ, Lith. *dubùs*, Goth. *diups* (Eng. *deep*), Toch. B *tapre* ‘high’

PIE \*b is reconstructed because it reflects as *b* in most languages. However, in Germanic, Armenian, Tocharian, and Anatolian, the reflex is *p*. In Germanic and Armenian this change is just a part of a larger chain shift (see below), while in Tocharian all stops merge into voiceless ones. The exact phonetics of Hitt. *p-*, *-p-* < PIE \*b (as opposed to *p-*, *-pp-* < PIE \*-p-) is not clear (p. 177).

There are not a lot of words containing PIE \*b. It seems to have been a marginal phoneme in PIE (it might have been completely absent in pre-PIE). The roots containing

PIE \*b are often attested in only a few languages, and sometimes this \*b seems to be of secondary origin. Clear cases of secondary \*b are found in forms where it originates from the \*ph<sub>3</sub> cluster – cf. the PIE root \*peh<sub>3</sub>- ‘drink’ (Ved. *apāt* ‘he drank’, Lat. *pōtus* ‘drunk’), where after reduplication one gets \*pi-ph<sub>3</sub>- > \*pib- ‘drink’ (Ved. *pībati* ‘drinks’, Lat. *bibō* ‘I drink’ < \*pibō, OIr. *ibid* ‘drinks’). In some cases, it seems likely that a PIE \*b originates in older \*m. Thus, the root \*h<sub>2</sub>eb1- ‘apple’ can be related to another PIE word for ‘apple’ – \*meh<sub>2</sub>lom (Lat. *mālum*, Gr. μῆλον, Alb. *mollë*; (?) Hitt. *māhlaš* ‘branch of a grapevine’). Here, one can assume a metathesis \*meh<sub>2</sub>l- > \*h<sub>2</sub>eml- and then the change \*-ml- > \*-bl- in northern IE dialects, where the word is attested. This kind of change can probably be seen in PIE \*bel- ‘strong, better’ (see above) as well – cf. Lat. *melior* ‘better’ < PIE \*mel-. A secondary \*b (presumably from older \*p) can perhaps be seen in PIE \*bol- (Lat. *blatea* ‘swamp’, OCS *blato* ‘mud’, Lith. *balà* ‘mud’, OEng. *pōl* > Eng. *pool*, Alb. *balë* ‘mud’), compared to PIE \*pol- (Skr. *palvalám* ‘pool, pond’, Lat. *palus* ‘swamp’, Lith. *pālios* ‘mud’). In some cases, there is even a variation of \*p/b/b<sup>h</sup>, cf. \*d<sup>h</sup>ewp- (PSl. \*dupa ‘hole’, ONor. *dúfa* ‘dive’, Eng. *dive* – but see p. 54), \*d<sup>h</sup>ewb- (see above), and \*b<sup>h</sup>ud<sup>h</sup>- ‘bottom’ (Ved. *budhnás*, Gr. πυθμήν, Lat. *fundus*, OEng. *bodan*) with metathesis. See p. 18 for more on PIE \*b.

### \*d

**PIE \*h<sub>1</sub>don(ts)** ‘tooth’ (< ‘eating’) > Ved. *dán* (Av. *dan̄tan-*), Gr. ὀδούς, Lat. *dēns* (⇒ Eng. *dentist*), PSl. \*dęsni ‘gums’ (BCMS *děsni*), Lith. *dantis*, Goth. *tunþus* (Eng. *tooth*), OIr. *dét*, Hitt. *adant-* ‘eating, eaten’, Arm. *atamn*

**PIE \*deks-** ‘right (side)’ > Ved. *dákšinas* (YAv. *dašina-*), Gr. δεξιός, Lat. *dexter* (⇒ Eng. *dexterity*), OCS *desnъ*, Lith. arch. *dėšinas*, Goth. *taihswa*, OIr. *dess* (Gaul. *Dex(s)iua* (name of a deity)), Alb. *djathtë*

**PIE \*doru** ‘tree, wood’ > Ved. *dāru* ‘wood’ (YAv. *dāru* ‘tree trunk’), Gr. δόρυ ‘tree’, OCS *drěvo* ‘tree’, Lith. *dervà* ‘resin, tar’, Goth. *triu* ‘tree trunk’ (Eng. *tree*), OIr. *daur* ‘oak’, Hitt. *tāru* ‘wood’, Alb. *dru*

PIE \*d is reconstructed because it reflects as *d* in most languages. The exceptions are again Germanic, Armenian, and Anatolian, with *t*, and Tocharian, with *ts* (cf. Toch. *tsām-* ‘grow’, Gr. δέμω ‘I build’ < PIE \*demh<sub>2</sub>-, p. 454).

### \*g

**PIE \*gerh<sub>2</sub>n/w-** ‘crane’ > Gr. γέρανος, Lat. *grūs*, PSl. \*žeravъ (OCz. *žeráv*) (with palatalization), Lith. *gėrvė*, OEng. *cran* (Eng. *crane*), MW/MBret. *garan* (Gaul. *tri-garanos* ‘with three cranes’), Arm. *krunk*

**PIE \*gem-** ‘squeeze, (com)press’ > Gr. γέμω ‘I am full/laden’, Lat. *gemō* ‘I groan’ (< \*‘groan under pressure/a load’), CS *gostъ* ‘dense’, Lith. *gāmalas* ‘lump’, (?) OIc. *kum(b)la* ‘to bruise, wound’, Mlr. *gemel* ‘shackle’, Toch. AB *kāmā-*

**PIE \*gel-/gl-ut-** ‘swallow; throat’ > Skr. *galas* ‘throat, neck’, Lat. *gluttō* ‘glutton’ (⇒ Eng. *gluttony*), ORuss. *gl̄tati* ‘to swallow’, OEng. *ceole* ‘throat’, (?) OIr. *geilid* ‘grazes, eats’, Arm. *klanem* ‘I swallow’

The phoneme \*g was, strangely, rather rare in PIE. It reflects as *g* in most languages (see p. 28 for palatalizations). The exceptions, with the voiceless reflex *k*, are, as usual, Germanic, Armenian, Anatolian, and Tocharian.

*PIE \*b and glottalic theory*

The fact that (early) PIE seems not to have had a phoneme \*b is typologically unusual, since this kind of gap (a language with *p* but no *b*) seems to be practically unattested elsewhere<sup>1</sup> (languages having a *b* and no *p*, like Arabic, do exist – Proto-Celtic was such a language as well). Together with the curious constraint in PIE that a root could not have two voiced stops (p. 54), this has led some researchers (Gamkrelidze & Ivanov 1995 being the most famous) to believe that PIE voiced stops ((\*b), \*d, \*g) should be phonetically reinterpreted as glottalized (ejective) stops ((\*p), \*t, \*k) (while the traditional \*p, \*t, \*k and \*b<sup>h</sup>, \*d<sup>h</sup>, \*g<sup>h</sup> would be reinterpreted as \*p<sup>(h)</sup>, \*t<sup>(h)</sup>, \*k<sup>(h)</sup> and \*b<sup>(h)</sup>, \*d<sup>(h)</sup>, \*g<sup>(h)</sup> – cf. p. 20–21). Since some languages with ejectives (like Navajo) are missing the bilabial stop *p* (though this is not the case in the majority of such languages – Quechua and Georgian have it, for instance), and sometimes have a constraint on having two ejectives in one word (though there are plenty of languages that do not), this could explain why (early) PIE had no \*b and no \*\*deg-type roots. The problem with this theory is that the reflexes of traditionally reconstructed PIE \*b, \*d, \*g are indeed voiced stops in the majority of IE languages/dialects. In cases where this is not true, the voiceless reflexes are hardly relevant – Germanic and Armenian have gone through complete consonantal shifts, and in Tocharian all stops became voiceless. Assuming an independent nontrivial change of glottalized voiceless to voiced stops in almost all IE branches is highly implausible, while the empirical evidence in favor of reconstructing PIE glottalized stops (cf. Kortlandt 1985) is interesting but hardly compelling. Thus, it seems best to stick with the traditional reconstruction of voiced (\*b), \*d, \*g for the last phase of PIE, though it is possible that pre-PIE indeed had glottalized \*t, \*k. Cf. also p. 154.

**Voiced aspirated stops**

The opposition of voiced and voiced aspirated stops is preserved in Indo-Aryan, in Italic (through fricativization of aspirates – p. 326–327), and, via complete consonant shifts, in Germanic and Armenian (p. 392–393, 426). Elsewhere (Anatolian, Celtic, Balto-Slavic, Albanian, Iranian), the aspiration is lost and the two series merge (traces of the old distinction are seen indirectly through Winter’s Law in Balto-Slavic – p. 481).

**\*b<sup>h</sup>**

**PIE \*b<sup>h</sup>reh<sub>2</sub>tēr** ‘brother’ > Ved. *bhrātā* (Av. *brātar-*), Gr. φράτηρ ‘member of a brotherhood’, Lat. *frāter* (⇒ Eng. *fraternity*), OCS *bratrъ*, Lith. *broterėlis* ‘little brother’, Goth. *broþar* (Eng. *brother*), OIr. *bráthir* (Gaul. *Bratronos* (pers. name)), Arm. *elbayr*, Toch. B *procer*

**PIE \*b<sup>h</sup>ewH-** ‘be, become, grow’ > Ved. *bhāvati* ‘is, becomes’ (GAv. *bavaŋ*), Gr. φύονται ‘they grow’, Lat. *fuī* ‘I were’, OCS *byti* ‘to be’, Lith. *būti* ‘to be’, Goth. *bauan* ‘to dwell, inhabit’ (Eng. *be*), OIr. *biid* ‘is wont to be’, Arm. *busanim* ‘I grow’

**PIE \*ǵomb<sup>h</sup>os** ‘tooth’ > Ved. *jám̐bhas*, Gr. γόμφος, OCS *zǫbъ*, Lith. *žam̐bas* ‘corner, edge’, ONor. *kambr* ‘comb’ (Eng. *comb*), Alb. *dhëmb*

In most branches (Germanic, Armenian, Iranian, Celtic, Balto-Slavic, Albanian), the reflex is a plain voiced *b* (with a merger with the reflex of PIE \*b, except in Germanic and Armenian, where PIE \*b > *p*). Anatolian *p-*, *-p-* has the same reflex as PIE \*b (cf. PIE \**-p-* > *-pp-*). Indic has voice together with aspiration (*bh*), while the reflex in Greek is a

voiceless aspirate  $\varphi$  [ $p^h$ ]. The aspiration is indirectly seen in Italic, where the reflexes are fricative, cf. Latin initial voiceless  $f$ - and medial voiced  $-b$ - (Proto-Italic fricative  $*-\beta$ -, p. 326–327). In Tocharian PIE  $*p$ ,  $*b$ , and  $*b^h$  merge into  $p$ . All this points to a PIE voiced aspirate (or breathy-voiced; p. 20–21)  $*b^h$ , with most languages preserving voice and some aspiration.

### $*d^h$

**PIE  $*h_1wid^hew$ -** ‘widow’ (perhaps originally: ‘unmarried’) > Ved. *vidhāvā* (YAv. *vidava*), Gr. ἡίθεος ‘bachelor’, Lat. *uidua*, OCS *vьdova*, OPruss. *widdewu*, Goth. *widuwo* (Eng. *widow*), OIr. *fedb*, Hitt. *w(i)dati*

**PIE  $*h_1rud^hros/h_1rowd^hos$**  ‘red’ > Ved. *rudhirás*, Gr. ἐρυθρός (⇒ Eng. *erythrocyte*), Lat. *ruber*, OCS *rьdrь*, Lith. *raūdas* ‘reddish’, OEng. *rēad* (Eng. *red*), OIr. *riúad* (Gaul. *Roudius* (pers. name)), Toch. B *ratre*

**PIE  $*d^huh_2mos$**  ‘smoke’ > Ved. *dhūmás*, Gr. θῦμός ‘soul, spirit’, Lat. *fūmus* (⇒ Eng. *fume*), OCS *dymь*, Lith. *dūmas*, OHG *doum*, MIr. *dumacha* ‘fog’, Hitt. *tuhhui-*, Toch. B *tweye* ‘dust, ashes’

Again, Indic has a voiced aspirated  $dh$  and Greek a voiceless aspirated  $\theta$  (written also  $\theta$ ) [ $t^h$ ]. In Balto-Slavic, Celtic, and Albanian, PIE  $*d^h$  >  $d$ , which merges with PIE  $*d$ . The merger is also seen in Hittite initial  $t$ - and medial (non-geminate)  $-t$ - (written also as  $<d>$ ). Germanic and Armenian (cf. p. 427) have voiced  $d$  with a consonantal shift, and in Tocharian PIE  $*d^h$  and  $*t$  merge into  $t$ . Proto-Italic has initial  $*\beta$ - and medial  $*-\delta$ -, reflected in Latin as  $f$ - and  $-d$ - (in most conditions) or  $-b$ - (before/after  $r$ , after  $u$ , and before  $l$ ).

### $*g^h$

**PIE  $*steyg^h$ -** ‘go, climb’ > Ved. *ati-ṣṭīgham* ‘to climb up’, Gr. στείχω ‘I walk’, OCS *po-stignōti* ‘to attain, to catch up with’, Lat. *stēigtiēs* ‘to hurry’, Goth. *steigan* ‘to climb’, OIr. *-tiagait* ‘walks’, Alb. *shteg* ‘footpath’

**PIE  $*h_3meyg^h$ -,  $*h_3mig^hle_h_2$**  ‘fog, cloud’ > Ved. *meghás* ‘cloud, rain’ (YAv. *maēya-*), Gr. ὀμίχλη ‘mist, fog’, OCS *mьgla* ‘fog’, Lith. *miqlà* ‘fog’, Dutch *miggelen* ‘to drizzle’, Arm. *mēg* ‘fog’, Alb. *mjegull* ‘fog, mist’

**PIE  $*g^h(e)h_2d^h$ -** ‘smooth’ > Lat. *glaber*, OCS *gladьkь*, Lith. *glodūs*, OHG *glat* ‘clear’ (Eng. *glad*)

PIE  $*g^h$  yields  $gh$  in Indic and  $\chi$  [ $k^h$ ] in Greek. In Iranian, Balto-Slavic, Celtic, and Albanian PIE  $*g^h$  (together with PIE  $*g$ ) yields  $g$ , the same as in Germanic and Armenian but in the latter two as a part of a general sound shift. In Hittite one finds  $k$ -,  $-k$ - (written also as  $<g>$ ), and in Tocharian all velars merge into  $k$ . In Latin we find  $h$  in most cases (the exceptions being *gr-/gl-*, *fu-*, and *-ng-*).

### Grassmann’s Law

In Indo-Aryan and Greek, there is an independent dissimilation rule that causes the deaspiration of the first aspirate in a word:  $C^h-C^h$  >  $C-C^h$ . Cf. PIE  $*b^hewd^h$ - > Ved. *bódhati* ‘is aware’ (not  $**bhódhati$ ), Gr. πρῶθουμαι ‘I give notice’ (not  $**φρῶθουμαι$ ). The fact that Greek exhibits voiceless deaspirated stops proves that the deaspiration occurred after the



Greek devoicing of PIE aspirated stops (which have a voiced reflex in Indic). This means that the deaspiration in Greek is younger and independent of Indic.

### *Consonantal shifts in Germanic and Armenian*

In Germanic and Armenian, all stops experience a very similar chain shift, in which all PIE stops weaken. Voiced aspirates lose aspiration, voiced stops lose voicedness, and voiceless stops change to fricatives (in Germanic) or to voiceless aspirates (in Armenian):

PIE	Germanic	Armenian
*b <sup>h</sup> , *d <sup>h</sup> , *g <sup>h</sup>	<i>b, d, g</i>	
*b, *d, *g	<i>p, t, k</i>	
*p, *t, *k	<i>f, þ, h</i>	*p <sup>c</sup> , t <sup>c</sup> , k <sup>c</sup>

Palatovelars and labiovelars change as well (see below). This seems to be an independent change in these two branches. In Germanic it is traditionally called **Grimm's Law** (p. 392–393, cf. p. 426 for Armenian). In Germanic PIE voiceless stops have a voiceless reflex if preceded by a fricative (*s, f, h, þ*), cf. PIE \*nok<sup>wts</sup> ‘night’ (Lat. *nox*) > Goth. *nahts* (not \*\*nahþs) or PIE \*pisk- (Lat. *piscis*) > OEng. *fisc* (Eng. *fish*).

In Germanic, Grimm's Law operates always in initial position. However, in the middle of the word it does so only if the syllable preceding PIE \*p/t/k was stressed in pre-Proto-Germanic (pre-PGmc had a free stress, like Vedic or Greek, which it subsequently lost – p. 396). If not (i.e., if the stress was on the syllable following \*p/t/k or elsewhere), **Verner's Law** (p. 393) operated, by which PGmc \*f, \*þ, \*x changed to voiced fricatives \*β, \*ð, \*γ (similar to Spanish <b, d, g> in most positions). In Gothic these yield *b, d, g*.

PGmc \*βrōþēr ‘brother’ (Ved. *bhrátā*, Gr. φράτηρ) > Goth. *broþar*

PGmc \*faðēr ‘father’ (Ved. *pitā*, Gr. πατήρ) > Goth. *faðar*

Grimm's and Verner's Laws operate on PIE palato- and labiovelars as well (p. 21–25). In Old High German (and Modern Standard German), there is another chain shift, called the OHG sound shift, where Germanic stops and fricatives change once again (p. 393–394).

### *The problem of the PIE voiced aspirates*

PIE voiced aspirates are reconstructed as such because this is what the reflexes point to – to aspiration or traces of it in some languages, and to voicedness in most of them. Indic is thought of as the most archaic since it preserves both voice and aspiration, though what are usually called “voiced aspirates” seem actually to be murmured/breathy-voiced stops (acoustically similar to aspirates) in many modern Indo-Aryan languages. Greek also has aspirates (though voiceless), and fricative reflexes in Italic (p. 326–327) point to original aspirates as well. In other branches, aspiration is lost, but the reflexes are always voiced (except in Tocharian, which is irrelevant). In Anatolian, where the exact pronunciation of stops is not clear (p. 177), one can still see that the reflexes of voiced and voiced aspirate stops are the same (non-geminate), while different from voiceless stops (geminate medially). Even in Germanic and Armenian, with a complete chain shift of stops, the internal logic of the shift (with each series “weakening” by one step – losing aspiration, voicedness, and plosiveness respectively) points to voiced aspirates. The typological parallel for

the development of PIE voiced aspirates can be found in Indo-Aryan, where Old Indic voiced aspirates yield voiced aspirates (breathy-voiced), voiceless aspirates, or voiced stops in various modern Indo-Aryan languages.

Thus, the reconstruction of PIE \*b<sup>h</sup>, \*d<sup>h</sup>, \*g<sup>h</sup> based on its reflexes is quite solid. However, the problem is again of typological nature – the PIE system of \*T, \*D, \*D<sup>h</sup>, i.e., the system with voiceless, voiced, and voiced aspirated stops (and no voiceless aspirates), is very unusual. It is a near universal in world languages that a language has to have voiceless aspirates if it has voiced aspirates (like in Indo-Aryan languages), or that, if it has just one set of aspirates, those are voiceless (like in Ancient Greek) – there is no reason for the double markedness of voiced aspirates (voice and aspiration) if voiceless aspirates are lacking in a language. However, as we have seen, the reflexes in IE languages do indeed point to voiced aspirates. Though the problem can hardly be considered solved, it may just be that PIE, with its \*T/D/D<sup>h</sup> system, was an (unstable) typological exception. And it is not the only one: there is at least one similar system in the world (with voiced aspirates but no voiceless aspirates) – in a dialect of the Austronesian Kelabit language.<sup>2</sup> Cf. also p. 18 and p. 153–154.

There is no real compelling evidence for the reconstruction of PIE voiceless aspirated stops, although they used to be reconstructed and sometimes still are by certain authors/schools. Indic voiceless aspirates originated in the combination of PIE \*p/t/k and laryngeals (but not \*ph<sub>3</sub> – p. 17), cf. PIE gen. sg. \*p<sub>nt</sub>Hos ‘path’ > Ved. gen. sg. *pathás*, Gr. *πάτος* (p. 77), \*h<sub>3</sub>esth<sub>1</sub> (p. 30), etc. For the laryngeal-aspiration of voiced plosives in Indo-Aryan cf. p. 30.

### Palatalized velars

Concerning the reflexes of PIE \*k, \*ǵ, \*ǵ<sup>h</sup>, which were palatalized velars (IPA [kʲ], [gʲ], [gʲ<sup>h</sup>]) rather than palatal stops (IPA [c], [j], [j<sup>h</sup>]), the IE languages can be divided into two groups – *centum* languages (after Lat. *centum* ‘hundred’ < PIE \*(d)k<sub>nt</sub>om) and *satem* languages (after Av. *satəm* ‘hundred’). In the *centum* group, PIE palatalized velars depalatalize, i.e., reflect in the same way as the original PIE \*k, \*g, \*g<sup>h</sup>. Most of the *centum* languages/branches are positioned to the west – Greek, Italic, Celtic, Germanic, and some fragmentarily attested languages (like Lusitanian and Venetic). The only exception is Tocharian, positioned to the east. In the *satem* group, PIE palatovelars usually reflect as either affricates (like *tš*, *tṣ̌*, *dž*, *dẓ̌* . . .) or fricatives (like *s*, *š*, *z*, *ž*, *b*, *ð* . . .), the fricatives usually being the further reflexes of older affricates (through processes like *tṣ̌* > *š*, *dẓ̌* > *ž*, etc.). Meanwhile, Balto-Slavic and Indo-Iranian (and fragmentarily attested Thracian, for instance) as *satem* languages are positioned more to the east. This is also true of Albanian, Armenian, and a part of Anatolian (Luwian and Lycian), which can be considered *satem* because of their reflexes of palatalized velars (but see p. 176). Considering that we find velars as reflexes in one group of languages and mostly affricates/fricatives (not depending on specific vowels, etc.) in other languages, the reconstruction of palatalized velars for the proto-language seems to be the most reasonable solution. For more on *centum/satem*, cf. p. 27–28.

### \*k

PIE \*k<sub>er</sub>(d), \*k<sub>rd</sub>- ‘heart’ > Ved. *śrad-dhā́* ‘faith’, Gr. κῆρ, καρδίᾱ/κρᾱδίᾱ (⇒ Eng. *cardiology*), Lat. *cor* (⇒ Eng. *cordial*), OCS *srъdьce*, Lith. *širdis* (Latv. *sirds*, OPruss. *seyr*), Goth. *hairto* (Eng. *heart*), OIr. *críde*, Hitt. *ker*, *kard(i)-* (Pal. *kārt-*, CLuw. *zārt-*, HLuw. *zart(i)-*), Arm. *sirt*, Toch. B *kāryāñ* ‘hearts’

**PIE \*h<sub>1</sub>ekwos** ‘horse’ > Ved. *ásvas* (YAv. *aspa-*), Gr. ἵππος (⇒ Eng. *hippo-drome*, *hippo(-potamus)*, Myc. *i-ǵo*), Lat. *equus* (⇒ Eng. *equestrian*), OLith. *ešva* ‘mare’, Goth. *aihwatundi-* ‘thornbush, bramble’ (lit. ‘horse-tooth’, OEng. *eoh* ‘warhorse’), OIr. *ech* (< Ogham Irish EQO-DDI, Clb. *Ekua-laku* (pers. name), Gaul. *Equos* ‘the ninth month’, *Epos* (pers. name)), HLuw. *azu-* (Lyc. *esb-*), Arm. *ēš* ‘donkey’, Toch. B *yakwe*, Ven. acc. sg. *ekvon*

**PIE \*h<sub>1</sub>nek-** ‘get, reach’ > Ved. *nás-* ‘attain’ (GAv. *nasaṭ* ‘reaches’), Gr. ἐνεγκεῖν ‘to carry’, Lat. *nancīscor* ‘I get’, OCS *nesti* ‘to carry’, Lith. *nėšti* ‘to carry’ (Latv. *nest*), Goth. *ga-nah* ‘to suffice’ (Eng. *enough*), OIr. *ro-ic* ‘reaches’, Arm. *hasanem* ‘I arrive’, Toch. B *enik-* ‘take, grip’

In *centum* languages, the reflexes are the same as for PIE \*k (p. 16). In *satem* languages, we find affricates and fricatives – OInd. *ś*, Av. *s*, OPers. *θ* (cf. OPers. *viθ-* ‘house(hold)’ to Ved. *viś-* ‘settlement’, Lat. *uīcus* ‘village’ < PIE \*wik-/woyk-) (p. 206), Slav./Latv./OPruss. *s*, Lith. *š* (< PBSL. \*ś), Alb. *th* [θ] (cf. Alb. *them* ‘I say’, Ved. *śamsati* ‘praises’, Lat. *cēnseō* ‘I suggest, propose’ < PIE \*ken/ms-), and Arm. *s* (PIE \*k > *š*, p. 434). In Luwian we find *z* [ts], and in Lycian *s* from PIE \*k before \*ē/i/y/w (p. 176, cf. always *k* in Hittite).

### \*ǵ

**PIE \*ǵe/onu** ‘knee’ (p. 74) > Ved. *jānu* (YAv. *zānu-*), Gr. γόνυ, Lat. *genū*, Goth. *kniu* (Eng. *knee*), Hitt. *gēnu*, Arm. *cunr*, Toch. A *kanweṃ* ‘knees’ (du.)

**PIE \*h<sub>2</sub>egros** ‘field’ > Ved. *ájras*, Gr. ἀγρός (⇒ Eng. *agro-nomy*), Lat. *ager* (⇒ Eng. *agri-culture*), Goth. *akrs* (Eng. *acre*)

**PIE \*ǵerh<sub>2</sub>-** ‘ripen, mature, old’, \*ǵr<sub>h</sub>nom ‘grain, corn’ > Ved. *járatī* ‘grows old’, *jīrnás* ‘old’ (YAv. *zar<sup>o</sup>ta-* ‘decrepit’), Gr. γέρων ‘old man’ (⇒ Eng. *ger-iatics*), Lat. *grānum* ‘grain’ (⇒ Eng. *grain*), OCS *zrělbъ* ‘ripe’, *zrěno* ‘grain’, Lith. *žirnis* ‘grain’ (Latv. *ziņnis*, OPruss. *syrne*), Goth. *kaurn* (Eng. *corn*), OIr. *grán* ‘grain’, Arm. *cer* ‘old’

In *centum* languages, the reflexes are the same as for PIE \*g (p. 17). In *satem* languages, we find mostly affricates and fricatives – OInd. *j* [dʒ], Av. *z*, OPers. *d* with depalatalization (cf. OPers. *dauštar-* ‘friend’, YAv. *zaoša-* ‘pleasure’, Ved. *jós-* ‘like’, Lat. *gustō* ‘I taste’ < PIE \*ǵews- ‘taste, like, choose’) – p. 206, Slav./Latv. *z* (written as <s> in OPruss.), Lith. *ž* (< PBSL. \*ž), Alb. *dh* [ð] (cf. Alb. *dhëmb* ‘tooth’ < PIE \*ǵomb<sup>h</sup>os, 567), Arm. *c* [ts] (with regular devoicing as with non-palatalized stops, p. 426).

### \*ǵ<sup>h</sup>

**PIE \*ǵ<sup>h</sup>eym-** ‘winter’ > Ved. *hēman* ‘in winter’ (cf. *himás* ‘cold, snow’ > *Himā-laya*; Av. *zəm-*, Pers. *dai*), Gr. χειμών, Lat. *hiems*, OCS *zima*, Lith. *žiema* (Latv. *ziema*, OPruss. *semo*), ONor. *gói* ‘month from the middle of February to the middle of March’, OIr. *gam* (Gaul. *Giamos* (pers. name)), Hitt. *gimm(ant)-*, Alb. *dimër*, Arm. *jivn* ‘snow’, Toch. A *šärme*

**PIE \*ǵ<sup>h</sup>ew-** ‘call’ > Ved. *hávate* ‘invokes’ (YAv. *zavati* ‘calls’), Gr. καυ-χάομαι ‘I speak loudly, boast’, OCS *zvati*, Lith. *žavėti* ‘charm, fascinate’, (?) Goth. *gub* ‘god’ (lit. \*‘the one who is called’, Eng. *god*), OIr. *guth* ‘voice’ (Gaul. *gutu-ater* ‘a

kind of druid’ < \*‘father of invocations’), Arm. *jaunem* ‘I devote’, Toch. B *kwā* ‘call out to, invite’

**PIE \*h<sub>2</sub>em/ng<sup>h</sup>**- ‘bind, squeeze; narrow’ > Ved. *ámh-* ‘compress’, *ámhas* ‘trouble’ (YAv. *qzah-* ‘difficulty’), Gr. ἄγγω ‘I squeeze’, Lat. *angō* ‘I squeeze/choke’, *angustus* ‘narrow’ (⇒ Eng. *anguish*), OCS *qzъkъ* ‘narrow’, Lith. *aĩkštas* ‘narrow’ (with *k* instead of *ž*), Goth. *aggwus* [ɲg] ‘narrow’ (Germ. *Angst* ⇒ Eng. *angst*, ONor. *angr* ⇒ Eng. *anger*), OIr. *cum-ung* ‘narrow’ (MW *ing*), Hitt. *hamank-* ‘tie’, Arm. *anjuk* ‘narrow’, Toch. B *entse* ‘greed, envy’ < \*enkse

In *centum* languages, the reflexes are the same as for PIE \*g<sup>h</sup> (p. 19). In *satem* languages, we find mostly affricates and fricatives – OInd. *h* (from older \*jh), Av. *z*, OPers. *d*, Slav./Latv. *z* (OPruss. <s>), Lith. *ž*, Alb. initially *d-* (cf. Alb. *dorë*, Gr. χεῖρ, Arm. *jeŋ* < PIE \*ǵ<sup>h</sup>esr- ‘hand’) or *dh-*, medially mostly *-dh-* (p. 567–568). In Armenian we find *j* [dʒ], a voiced pair of *c* < PIE \*ǵ. In languages where the reflex of PIE \*g is the same as the reflex of \*g<sup>h</sup>, PIE \*ǵ<sup>h</sup> has the same reflex as \*ǵ, the only exception being Albanian.

TABLE 1.2 THE REFLECTION OF PIE PALATOVELARS IN *SATEM* LANGUAGES

PIE	Luw.	Lyc.	OInd.	Av.	OPers.	Lith.	Latv.	Slav.	Arm.	Alb.
*k	<i>z</i> [ts]/ <i>k</i>	<i>s</i> / <i>k</i>	<i>ś</i>	<i>s</i>	<i>θ</i>	<i>š</i>	<i>s</i>	<i>s</i>	<i>s</i>	<i>th</i> [θ]
*ǵ	<i>(k/y/∅)</i>		<i>j</i> [dʒ]	<i>z</i>	<i>d</i>	<i>ž</i>	<i>z</i>	<i>z</i>	<i>c</i>	<i>dh</i> [ð]
*ǵ <sup>h</sup>	<i>(k/y/∅)</i>		<i>h</i>	<i>z</i>	<i>d</i>	<i>ž</i>	<i>z</i>	<i>z</i>	<i>j</i> [dʒ]	<i>d/dh-</i> , <i>-dh-</i>

### Labiovelars

The *centum/satem* distinction is relevant concerning the PIE labialized velars (\*k<sup>w</sup>, \*g<sup>w</sup>, \*g<sup>w</sup>h) as well. In the *centum* group (where \*K̄ = \*K), they usually yield different reflexes than the regular velars (\*k, \*g, \*g<sup>h</sup>). The labiovelars either are preserved (as *k<sup>w</sup>*, *g<sup>w</sup>*, *x<sup>w</sup>*, etc.) or are (subsequently) changed to full labials (*p*, *b*, *w*) or plain velars (*k*, *g*). Tocharian seems to be the only *centum* language that shows no distinction between old plain and labialized velars (p. 454). In the *satem* group (where \*K̄ ≠ \*K), the labiovelars usually reflect as plain velars (\*K = \*K<sup>w</sup>) – this is the case in Balto-Slavic and Indo-Iranian (although there are perhaps some indirect traces of labiovelars in Indo-Iranian). But in Albanian (p. 567) the reflexes of plain and labialized velars are distinct before front vowels (\*K ≠ \*K<sup>w</sup>), and the same is true for \*k ≠ \*k<sup>w</sup> before front vowels in Armenian (p. 427). In Luwo-Lycian, which preserves a special reflex of \*k in certain conditions (p. 176), the PIE labiovelars have separate reflexes as usual in Anatolian. Traditionally, the *centum* group is represented by the formula \*k = \*k̄ ≠ \*k<sup>w</sup> and the *satem* group by \*k = \*k<sup>w</sup> ≠ \*k̄, but Armenian, Albanian, and Luwo-Lycian (and Tocharian with *k* from all PIE velars) cannot be neatly subsumed into these formulas (see p. 27).

### \*k<sup>w</sup>

**PIE \*k<sup>w</sup>etwores** ‘four’ > Ved. *catváras* (YAv. *čaθβārō*) (with palatalization), Gr. τέτταρες, Lat. *quattuor*, OCS *četyre* (with palatalization), Lith. *keturi*, Goth. *fidwor* (Eng. *four*), OIr. *ceth(a)ir* (Gaul. *petuar-*, OW *petguar*), (?) Alb. *katër* (if not from Lat.), Arm. *čork* (with palatalization), Toch. B *štwer* (with palatalization)

- PIE \*penk<sup>w</sup>e** ‘five’ > Ved. *pāñca* (YAv. *pañča*) (with palatalization), Gr. πέντε (⇒ Eng. *penta-gram*), Lat. *quīnque*, OCS *petь* < \*penktis, Lith. *penki*, Goth. *fimf* (Eng. *five*), OIr. *cóic* (Gaul. *pempe-*, OW *pimp*), Alb. *pesë* (with palatalization)
- PIE \*k<sup>w</sup>el-** ‘turn’, **\*k<sup>w</sup>ek<sup>w</sup>los** ‘wheel’ (with root reduplication) > Ved. *cakrá-* ‘wheel’ (⇒ Eng. *chakra*, YAv. *čaxra-* ‘wheel’) (with palatalization), Gr. κύκλος ‘ring, circle’ (⇒ Eng. *cycle*), πόλος (< \*k<sup>w</sup>olos) ‘axis’ (⇒ Eng. *pole*), τέλος (< \*k<sup>w</sup>elos) ‘end, outcome’ (< \*‘full circle’, ⇒ Eng. *teleo-logy*), Lat. *colus* ‘distaff’, OCS *kolo* ‘wheel, circle’, Lith. *kāklas* ‘neck’, OEng. *hwēol* (with Grimm’s Law, PGmc \*x<sup>w</sup>ex<sup>w</sup>l-, cf. Eng. *wheel*), *hweogul* (with Verner’s Law), MIr. *cul* ‘chariot’, Toch. B *kokale* ‘cart, wagon’, Alb. *sjell* ‘bring’ (with palatalization)

In *centum* languages, PIE \*k<sup>w</sup> remains unchanged in part of Italic (Lat. <qu> but Sabelian *p*), Anatolian (Hitt./Luw. <ku>), Mycenaean Greek (<q>), and some Celtic languages (Ogham Irish <Q>, CIB. <ku>). Proto-Greek \*k<sup>w</sup> changes to π, τ, or κ in Attic, depending on what comes next (πα, πο, πC, τι, τε, κν – p. 292). In Q-Celtic languages, Proto-Celtic \*k<sup>w</sup> remains unchanged or loses the labialization later (OIr. *c*), while in P-Celtic it changes to *p* (p. 357–358, 363). In Germanic the reflex is \*x<sup>w</sup> (Goth. <hu>, OEng. <hw>, Eng. <wh> – still pronounced [h<sup>w</sup>] in some dialects) by Grimm’s, or \*γ<sup>w</sup> by Verner’s Law (p. 20, 392). In certain cases, one finds \*f instead of \*x<sup>w</sup> in Germanic (like in ‘four’ and ‘five’ above, or ‘wolf’ – p. 35), which can be explained in different ways. Tocharian has *k* for all velars. In *satem* languages, PIE \*k<sup>w</sup> usually yields *k*, just like PIE \*k. However, Luwian preserves the distinction of *ku* < PIE \*k<sup>w</sup>, *k* < PIE \*k, and *z* [c] < PIE \*k̄ in certain conditions (p. 176). One finds different outcomes of PIE \*k<sup>w</sup>, \*k, and \*k̄ before front vowels (PIE \*e, \*i) in Albanian and Armenian as well – Alb. *s* : *k* : *th* and Arm. *č*<sup>o</sup> : *k*<sup>o</sup> : *s* (see p. 427, 567).

### The reflex of PIE \*kw/kw in *centum* languages

PIE \*k<sup>w</sup> seems to have been distinct from PIE \*kw (\*k + \*w, likewise \*g<sup>w(h)</sup> ≠ \*g<sup>(h)w</sup>), though the latter was rare. In Greek, PIE \*kw/kw > κκ (though this is disputed), cf. Gr. λάκκος ‘pond’, OCS *loky* ‘puddle’, Lat. *lacus* ‘lake’ < PIE \*laku/w- (or \*lok-/l̥k-). The reflex of PIE \*kw in Latin is not clear (perhaps it is *c* in initial position before *a*), cf. Lat. *equus* (p. 22) and *queror* ‘I lament’ (Ved. *śváṣiti* ‘blows’ < PIE \*k<sup>w</sup>es-, but *cāseus* ‘cheese’ (OCS *kvasь* ‘leaven’ < PIE \*k<sup>w</sup>eh<sub>23</sub>/ō(t)s-) and *canis* ‘dog’ (OInd. *śvā* < PIE \*k<sup>w</sup>on-). PIE \*kw/kw yields PCelt. \*k<sup>w</sup> (cf. PCelt. \*ek<sup>w</sup>os ‘horse’), but not in Germanic and Tocharian (cf. Goth. *aihwā-*, Toch. B *yakwe*, p. 22).

### \*g<sup>w</sup>

- PIE \*g<sup>w</sup>ōws** ‘cow’ > Ved. *gāus* (Av. *gāuš*), Gr. βοῦς, Lat. *bōs* (a Sabellian loanword, should be \*\*uōs in Lat.), OCS *govęzdb* ‘bovine’, Latv. *gūovs*, OEng. *cū* (Eng. *cow*), OIr. *bó* (Gaul. *Bo-marus* (pers. name)), HLuw. *wawa/i-* (Lyc. *wawa-*), Arm. *kov*, Toch. B *ke<sub>u</sub>*
- PIE \*g<sup>w</sup>en(e)h<sub>2</sub>** ‘woman’ > Ved. *jānī* (with palatalization), *gnās* ‘divine woman’ (GAV. *jāni-*, *ganā-*), Gr. γυνή (⇒ Eng. *gynē-cology*, dial. Boeotian βανᾶ), OCS *žena*, OPPruss. *genno*, Goth. *qino* (cf. Eng. *queen* with semantic change), OIr. *bé/ben* (Gaul. gen. pl. *bnanom*), CLuw. *wānā-*, Arm. *kin*, Toch. B *šana* (with palatalization), Messapic *benna*

**PIE \*g<sup>w</sup>e/orH-/\*g<sup>w</sup>rH-** ‘mountain’ > Ved. *giris* (YAv. *ga’ri-*), Gr. δειράς ‘ridge of a chain of hills’, βορέης ‘north wind’, OCS *gora* (Slav. also ‘forest’), Lith. *girià* ‘forest’, Alb. *gur* ‘stone’

In *centum* languages, PIE \*g<sup>w</sup> as such is not preserved anywhere (except partially in Latin after *-n-*). In Latin the usual reflex is *u* [w] (the exceptions being *gr-/gl-* and *-ngu-*, while in Sabellian one finds *b* – p. 327). Hittite has <ku>, and Luwian has *w*. Proto-Greek \*g<sup>w</sup> changes to β, δ, γ in Attic (βα, βo, βC, βi, δε, γv – p. 292), but not in complete parallelism to voiceless counterparts (thus βi < \*g<sup>w</sup>i, but τi < \*k<sup>w</sup>i). PIE \*g<sup>w</sup> yields PCelt. \*b (while PIE \*k<sup>w</sup> is preserved and PIE \*g<sup>wh</sup> > PCelt. \*g<sup>w</sup>, p. 363), and, according to Grimm’s Law, PGmc \*k<sup>w</sup> (Goth. <q>, OEng. <cw>/Eng. <qu>, OHG <qu>, ONor. <kv>). In *satem* languages (including Armenian but not Albanian before front vowels), PIE \*g<sup>w</sup> is reflected the same as PIE \*g.

### \*g<sup>wh</sup>

**PIE \*sneyg<sup>wh</sup>-** ‘snow’ > Ved. *snih-* ‘wet(ness)’ (‘snow’ in Prakrits, YAv. *snaēža-* ‘snow’, with palatalization), Gr. acc. sg. νίφα, Lat. *nix* (< \*nig-s), gen. sg. *niuis*, *ninguit* ‘it snows’ (with *-n-* infix, p. 95), OCS *sněgъ*, Lith. *sniēgas*, Goth. *snaiws* (Eng. *snow*), OIr. *snige*

**PIE \*g<sup>wh</sup>en-** ‘strike, kill’ > Ved. *ghnánti* ‘they strike, kill’ (< \*g<sup>wh</sup>enti), *hánti* ‘strikes, kills’ (< \*g<sup>wh</sup>enti, with palatalization) (YAv. *ja’nti* ‘strikes, kills’), Gr. φνεύω ‘I kill’, θείνω ‘I strike’, Lat. *dē-fendō* ‘I ward off, repel’ (⇒ Eng. *defend*), OCS *goniti* ‘to chase’, *žęti* ‘to reap’ (with palatalization), Lith. *giñti* ‘to drive (animals)’, ONor. *gunnr* ‘war, battle’, OIr. *gonaid* ‘strikes, kills’, Hitt. *kuenzi* ‘kills’, Arm. *gan* ‘beating’, Toch. B *käsk-* ‘scatter to destruction’

**PIE \*dheg<sup>wh</sup>-** ‘burn’ > Ved. *ni-dāghás* ‘heat’, *dáhati* ‘burns’ (YAv. *daža’ti* ‘burns’, with palatalization), Gr. τέφρα ‘ashes’, Lat. *febris* ‘fever’ (⇒ Eng. *fever*), *foueō* ‘I (keep) warm’, OCS *žęgъ* ‘I burn’ (< \*geg- < \*deg- with assimilation), Lith. *dęgti* ‘to burn’, Goth. *dags* (Eng. *day*), OIr. *daig* ‘flame’, Alb. *djeg*, Toch. AB *tsäk-* ‘burn up’

In *centum* languages, the labialization of PIE \*g<sup>wh</sup> is usually not preserved (except in Latin and Gothic after *-n-*). In Proto-Italic one expects \*h<sup>w-</sup>, \*-γ<sup>w-</sup>. In Latin this yields initial *f-* and medial *-u-* (but *-ngu-* and *-br-*). Proto-Greek \*k<sup>wh</sup> (with Greek devoicing of aspirates) changes to φ, θ, χ in Attic (φα, φο, φC, φi, θε, χv – p. 292), in parallelism to voiced counterparts. PIE \*g<sup>wh</sup> yields PCelt. \*g<sup>w</sup> (> OIr. *g*, p. 363). In Germanic, PIE \*g<sup>wh</sup> yields *w* or *g* (before *u*, *o*, and consonants), although not everything is clear. After *n*, the expected \*g<sup>w</sup> is seen (cf. Goth. *saggws* [sang<sup>w</sup>s] ‘song’, Gr. ὀμῆ < PIE \*song<sup>wh-</sup>). In some cases, *b* is also posited as a reflex by some; e.g., both Goth. *brinnan* (Eng. *burn*) and OEng. *wearm* (Eng. *warm*) are sometimes derived from PIE \*g<sup>wh</sup>erm- ‘burn’ (Gr. θερμός ‘hot’, Lat. *formus* ‘warm’, etc.), but alternative etymologies exist as well. In Anatolian, Balto-Slavic, and Albanian, PIE \*g<sup>wh</sup> reflects the same as \*g<sup>w</sup> (in Armenian, \*g<sup>wh</sup> reflects as \*g<sup>h</sup>). In Albanian, PIE \*g<sup>w(h)</sup> merges with PIE \*g<sup>(h)</sup> before back vowels (and yields *g*), while in front of \*e/i it yields *z* (cf. PIE \*g<sup>wh</sup>ermo- > Alb. *zjarr* ‘fire, fever’) unlike normal velars.

### Loss of labiality

In some *centum* languages that preserve PIE labiovelars in one way or another, there is a tendency toward loss of labiality in some positions – i.e., in front of rounded vowels (*u*

and sometimes *o* – a case of phonetic dissimilation). This explains Gr. *κν, γν, χν* (with plain velars before \**u*) and examples like Lat. *cultūra* ‘cultivation’, *incola* ‘inhabitant’, but *inquilīnus* ‘tenant’. In Latin the labiality is lost before consonants as well, cf. Lat. *reliquiae* ‘remains’ but *relictus* ‘left behind’.

### The problem of three velar series in PIE

According to the most usual and traditional reconstruction, dating from the 19th century and Karl Brugmann, PIE had three velar series (regular, palatalized, and labialized):

\**k*, \**g*, \**g<sup>h</sup>*  
 \**k̑*, \**ǵ*, \**ǵ<sup>h</sup>*  
 \**k<sup>w</sup>*, \**g<sup>w</sup>*, \**g<sup>wh</sup>*

This kind of system is complex and not typologically too frequent, but it is nonetheless attested – e.g., in the modern Iranian language Yazgulyam<sup>3</sup> (where it is an innovation of no historical relation to the PIE three series). There is a long tradition of denying the existence (or originality) of the three velar series in PIE, starting with Antoine Meillet and, surprisingly, still alive today. In this approach, the palatovelars are usually considered somehow secondary.

The main argument against the PIE distinction of \**K* : \**K̑* : \**K<sup>w</sup>* is that the majority of the languages distinguish only the reflexes \**K* : \**K̑* or \**K* : \**K<sup>w</sup>*. However, as already said, the traditional formulas (\**K* = \**K̑* ≠ \**K<sup>w</sup>* for *centum* and \**K* = \**K<sup>w</sup>* ≠ \**K̑* for *satem* languages) are not valid for all languages. Albanian (p. 567) and Armenian (p. 426–427) distinguish all three series before front vowels (Armenian only in voiceless stops), while Luwian and Lycian distinguish the reflexes of PIE \**k* : \**k̑* : \**k<sup>w</sup>* (in voiceless stops) in certain conditions (p. 176), cf. CLuw. *kiš-* ‘to comb’ < PIE \**kes-* (p. 16), CLuw. *zārt-* ‘heart’ < PIE \**k̑rd-* (p. 21), CLuw. *kui-* ‘who, what’ < PIE \**k<sup>wi-</sup>* (Lat. *quis*, p. 88).<sup>4</sup> The adversaries of the PIE three series usually offer no plausible alternate solutions for these, but simply hold, if only implicitly, that Anatolian evidence is scarce and inconclusive, or that Armenian and Albanian are too innovative and attested too late to be convincing.

A convincing argument for the three original series is found in the fact that PIE \**K* : \**K̑* : \**K<sup>w</sup>* cannot be predicted from phonemes in their surroundings (though some tendencies do exist). Thus, we find palatovelars in contact with front vowels (direct or indirect, before or after them): \**kerd-* (p. 21), \**ǵ<sup>h</sup>eym-* (p. 22), \**h<sub>1</sub>ekwos* (p. 22), \**ǵneh<sub>3-</sub>* ‘know’ (Ved. *jānāti* ‘knows’, OCS *znati* ‘to know’, Goth. *kann* ‘to get to know’, etc.), etc. However, we also find palatovelars when not in contact with front vowels: \**kwō(n)* (p. 41), \**klownis* (p. 51), \**kolh<sub>2</sub>m-/k̑lh<sub>2</sub>m-* ‘straw’ (Gr. *κάλαμος* ‘reed, stalk’, Lat. *culmus*, PSI. \**solma* > BCMS *slāma*, OEng. *healm* > Eng. *haulm*), etc. In cases like \**ǵomb<sup>h</sup>os* (p. 18), one could argue the connection with the verbal stem \**ǵemb<sup>h</sup>-*, etc. Likewise, regular velars are found even when front vowels follow – \**kes-* (p. 16), \**kelh<sub>3-</sub>* ‘rise’ (Lat. *ex-cellō* ‘I rise’, Lith. *kėlti* ‘to lift, raise’, etc.), \**gerh<sub>2</sub>n/w-* (p. 17), etc., though, of course, one can always assume analogies in some cases (to variants \**kos-*, \**kolh<sub>3-</sub>*). There are also pairs like \**leyǵ<sup>h</sup>-* (p. 32), \**steyǵ<sup>h</sup>-* (p. 19), \**sneyǵ<sup>wh</sup>-* ‘to snow’ (p. 25), and even “minimal pairs” like \**b<sup>h</sup>erǵ<sup>h</sup>-* ‘take care of’ (OCS *ne-brěgo* ‘I neglect’, Goth. *baīrgan* ‘to preserve, keep’) and \**b<sup>h</sup>erǵ<sup>h</sup>-* ‘tall, high’ (Ved. *bṛhānt-*, OHG *berg* ‘hill’, Arm. *barjr* ‘high’).<sup>5</sup> All this proves that palatovelars were phonemic in the last phase of PIE, and that they did

not depend synchronically on the supposed phonetic environment prone to palatalization. Certain possible PIE variants like *\*lew̥k-/lew̥k-* ‘shine’ (cf. Ved. *rokás* ‘light’ but *rúsant-* ‘shining’) are probably just remnants of some pre-PIE leveling, perhaps similar to PIE variants with *\*p/b/b<sup>h</sup>* (p. 17).

One usual argument against three series of velars in PIE is the unexpected frequency of these three series. One would expect that the unmarked plain velars would be the most frequent, while in reality they are the least frequent. However, this in no way proves that palatovelars are “secondary” or something of the sort. It just means that the system was unstable, which is clear from the fact that it was not preserved in any of the IE languages. The low frequency of plain velars is surely to be accounted for by certain conditions in the PIE prehistory. A frequent proposition is that PIE plain velars historically derive from uvular velars (thus pre-PIE *\*q* > PIE *\*k*, while, for instance, pre-PIE *\*k* > PIE *\*k<sup>6</sup>*). Another possible hypothesis is that the PIE velar system is to be derived from an older system with plain velars only and a number of vowels, similar to Proto-Uralic, which is often mentioned as the most likely relative of IE (p. 7). It is imaginable that a system like that, after a massive merger of vowels (p. 42–43), would result in what we have in PIE. If, let us suppose, PIE *\*k<sup>w</sup>(e)* is to be derived from pre-PIE *\*ku*, *\*ko*, PIE *\*k(e)* from pre-PIE *\*ki*, *\*ke* (and possibly from *\*kü* and *\*kā* as well), and PIE *\*k(e)* from pre-PIE *\*ka* only, it is logical to expect that PIE *\*k* would be the least frequent.

As for the so-called *centum* reflexes in *satem* languages, which are sometimes adduced as proof of the secondary/loose nature of the three series, these seem to be mostly of secondary, post-PIE nature – due either to later depalatalizations in certain *satem* languages (like the regular depalatalization before *\*m/l/r* and back vowel in Balto-Slavic, cf. Lith. *akmuō* (p. 41) but *āšmenys* ‘edge, blade’<sup>7</sup>), or loanwords from *centum* languages (like OCS *brěgъ* ‘bank, shore’, Slav. also ‘hill’ < PSI. *\*bergъ* probably from Germanic – see above for PIE *\*b<sup>h</sup>erǵ<sup>h</sup>-*). Some of the *centum* reflexes may also be explained by rare PIE palatalized/non-palatalized variants like *\*lew̥k-/lew̥k* (see above).

### **Centum and satem languages**

As already mentioned (p. 21, 23), the *centum/satem* distinction in IE is based on different treatment of PIE *\*Ķ* and *\*K<sup>w</sup>* in IE branches/languages. This used to be considered a major dialectal isogloss in PIE that divided PIE into western (*centum*) and eastern (*satem*) dialects. However, this view is nowadays untenable. The west/east divide was broken already by Tocharian (which is the easternmost IE language), but more important is the fact that the traditional formulas (*\*K* = *\*Ķ* ≠ *\*K<sup>w</sup>* for *centum* and *\*K* = *\*K<sup>w</sup>* ≠ *\*Ķ* for *satem* languages) do not work for Armenian, Albanian, and Anatolian. Although other interpretations are possible (since this is essentially a question of terminology, not essence), the *centum/satem* distinction could be explained as follows – in *satem* languages, PIE *\*Ķ* yields affricates/fricatives, while the labiality of PIE *\*K<sup>w</sup>* is usually (though not always and everywhere) lost. The *centum* group would then consist of those languages where PIE *\*Ķ* > *K* (while *\*K<sup>w</sup>* is usually, though not always and forever, preserved). Thus, Albanian, Armenian, and Luwian would be *satem* and Tocharian would be *centum*. The fact that in Anatolian one finds both *satem* (Luwian, Lycian) and *centum* languages (Hittite, Palaic), and the fact that often very close PIE dialects (like Indo-Iranian and Greek) are divided by this isogloss would point to the conclusion that this is not a crucial dividing point in PIE, as was once thought, but is a rather trivial phonetic innovation/variable.





centum languages  
 satem languages  
 centum/satem languages

MAP 1.1 THE CENTUM-SATEM BRANCHES OF IE

**Later palatalizations**

In many IE branches, sub-branches, and languages, later palatalizations of the PIE velars (and, less often, other phonemes like dentals) occur before front vowels – e.g., in Indo-Iranian (p. 226, 270), Slavic (p. 528–529), Latvian (p. 496), Germanic, Anatolian (e.g., Hitt. \**ti* > *zi*, \**di* > *ši*), Armenian (p. 427), Albanian (p. 569), Tocharian (p. 459–460), and most Romance languages (but not in Classical Latin). The change of \**Cy* > *C<sup>j</sup>*, which can also be interpreted as a kind of palatalization, can also be seen in, e.g., Lithuanian (p. 496), Slavic (p. 529–530), and Greek (p. 293).

TABLE 1.3 BASIC REFLEXES OF PIE STOPS (palatalizations and other special changes, like Verner’s Law, are not included)

PIE	OInd.	Av.	Gr.	Lat.	OCS	Lith.	Goth.	Hitt.	Luw.	Celt.	Arm.	Alb.	Toch.
*p	<i>p</i>	<i>p</i>	<i>π</i>	<i>p</i>	<i>p</i>	<i>p</i>	<i>f</i>	<i>p-</i> , <i>-pp-</i>	<i>p-</i> , <i>-pp-</i>	*∅	<i>h-/Ø-</i> , <i>-w-</i>	<i>p</i>	<i>p</i>
*t	<i>t</i>	<i>t</i>	<i>τ</i>	<i>t</i>	<i>t</i>	<i>t</i>	<i>þ</i>	<i>t-</i> , <i>-tt-</i>	<i>t-</i> , <i>-tt-</i>	*t	<i>tʰ</i>	<i>t</i>	<i>t</i>
*k	<i>k</i>	<i>k</i>	<i>κ</i>	<i>c</i>	<i>k</i>	<i>k</i>	<i>h</i>	<i>k-</i> , <i>-kk-</i>	<i>k-</i> , <i>-kk-</i>	*k	<i>kʰ</i>	<i>k</i>	<i>k</i>
*b	<i>b</i>	<i>b</i>	<i>β</i>	<i>b</i>	<i>b</i>	<i>b</i>	<i>p</i>	<i>p</i>	<i>p</i>	*b	<i>p</i>	<i>b</i>	<i>p</i>
*d	<i>d</i>	<i>d</i>	<i>δ</i>	<i>d</i>	<i>d</i>	<i>d</i>	<i>t</i>	<i>t</i>	<i>t</i>	*d	<i>t</i>	<i>d</i> , <i>-dh/d-</i>	<i>ts</i>
*g	<i>g</i>	<i>g</i>	<i>γ</i>	<i>g</i>	<i>g</i>	<i>g</i>	<i>k</i>	<i>k</i>	<i>k</i>	*g	<i>k</i>	<i>g</i>	<i>k</i>
*b <sup>h</sup>	<i>bh</i>	<i>b</i>	<i>φ</i>	<i>f-</i> , <i>-b-</i>	<i>b</i>	<i>b</i>	<i>b</i>	<i>p</i>	<i>p</i>	*b	<i>b</i> , <i>-w-</i>	<i>b</i>	<i>p</i>
*d <sup>h</sup>	<i>dh</i>	<i>d</i>	<i>ϑ</i>	<i>f-</i> , <i>-d/b-</i>	<i>d</i>	<i>d</i>	<i>d</i>	<i>t</i>	<i>t</i>	*d	<i>d</i>	<i>d</i>	<i>t</i>
*g <sup>h</sup>	<i>gh</i>	<i>g</i>	<i>χ</i>	<i>h</i>	<i>g</i>	<i>g</i>	<i>g</i>	<i>k</i>	<i>k</i>	*g	<i>g</i>	<i>g</i>	<i>k</i>
*k	<i>s</i>	<i>s</i>	<i>κ</i>	<i>c</i>	<i>s</i>	<i>š</i>	<i>h</i>	<i>kk</i>	<i>z</i> , <i>k</i>	*k	<i>s</i>	<i>th</i>	<i>k</i>
*ǵ	<i>j</i>	<i>z</i>	<i>γ</i>	<i>g</i>	<i>z</i>	<i>ž</i>	<i>k</i>	<i>k</i>	<i>k-</i> , <i>-Ø-</i>	*g	<i>c</i>	<i>dh</i>	<i>k</i>
*ǵ <sup>h</sup>	<i>h</i>	<i>z</i>	<i>χ</i>	<i>h</i>	<i>z</i>	<i>ž</i>	<i>g</i>	<i>k</i>	<i>k-</i> , <i>-Ø-</i>	*g	<i>j</i>	<i>d/dh-</i> , <i>-dh-</i>	<i>k</i>
*k <sup>w</sup>	<i>k</i>	<i>k</i>	<i>π</i> , <i>τ</i> , <i>κ</i>	<i>qu</i>	<i>k</i>	<i>k</i>	<i>hw</i>	<i>ku-</i> , <i>-kku-</i>	<i>ku-</i> , <i>-kku-</i>	*k <sup>w</sup>	<i>kʰ</i>	<i>k</i>	<i>k</i>
*g <sup>w</sup>	<i>g</i>	<i>g</i>	<i>β</i> , <i>δ</i> , <i>γ</i>	<i>u</i>	<i>g</i>	<i>g</i>	<i>q</i>	<i>ku</i>	<i>w</i>	*b	<i>k</i>	<i>g</i>	<i>k</i>
*g <sup>wh</sup>	<i>gh</i>	<i>g</i>	<i>φ</i> , <i>ϑ</i> , <i>χ</i>	<i>f-</i> , <i>-u-</i>	<i>g</i>	<i>g</i>	<i>w</i>	<i>ku</i>	<i>w</i>	*g <sup>w</sup>	<i>g</i>	<i>g</i>	<i>k</i>

## PIE FRICATIVES

PIE fricatives were \*s and the so-called laryngeals (\*h<sub>1</sub>, \*h<sub>2</sub>, \*h<sub>3</sub>), or at least \*h<sub>2</sub> and \*h<sub>3</sub>, which were almost certainly some kind of *h*-sounds. In PIE, \*z was a voiced allophone of \*s, like in \*nisdos [nizdos] ‘nest’ (OCS *gnězdo*, OEng. *nest*). For PIE allophones \*p, \*ð, \*ð<sup>h</sup> see p. 53.

In IE languages, PIE \*s either remains unchanged or changes (in various conditions) to *š*, *h* (glottal fricative), *x* (velar fricative), *z*, or (*z* > *ž* >) *r*. The change to *h*, typologically common (cf. many varieties of modern Spanish or many Polynesian languages), occurs independently in Iranian, Greek (initially and originally intervocalically), and Armenian (similar to Greek), and in some sandhi-conditions in Old Indic and Insular Celtic (cf. also Slavic, where PIE \*s > \*š > *x* in certain conditions – p. 526–527). The change to *r* (rhotacism), also typologically common (cf. the sporadic intervocalic *ž* > *r* in Western South Slavic, e.g., Sln. *móre* ‘can’ to OCS *možetb*), occurs independently in Latin, North-West Germanic, Eretrian Ionic Greek dialect, and in certain sandhi-positions in Old Indic.

## \*s

**PIE \*septm** ‘seven’ > Ved. *saptá* (YAv. *hapta*), Gr. *ἑπτὰ*, Lat. *septem* (⇒ Eng. *September*), OCS *sedmъ*, Lith. *septyni*, Goth. *sibun* (Eng. *seven*), OIr. *secht*<sup>r</sup> (Gaul. *sextan-*), Hitt. *šipta-miya-* ‘seven-drink’, Arm. *ewr<sup>r</sup>n*, Alb. *shtatë*, Toch. A *spät*

**PIE \*(sti-)steh<sub>2</sub>-** ‘stand’ > Ved. *sthā-*, *tīṣṭhati* ‘stands’ (YAv. *hišta’ti* ‘stands’), Gr. *ἵστημι* ‘I make to stand’, *στατός* ‘placed, standing’, Lat. *stō* ‘I stand’, *sistō* ‘I place (cause to stand)’, OCS *stati* ‘to stand’, Lith. *stóti* ‘to stand’, Goth. *standan* ‘to stand’ (Eng. *stand*), OIr. *sissidir* ‘stands’ (Cib. SISTAT ‘puts’), Hitt. *ištantāye-* ‘stay put’, Toch. B *ste* ‘is’

**PIE \*muHs** ‘mouse’ > Ved. *múš-*, Gr. *μῦς* – gen. sg. *μῦός*, Lat. *mūs* – gen. sg. *mūris*, OCS *myšb*, OEng. *mūs* (Eng. *mouse*)

Except in special cases, PIE \*s does not change in Anatolian (written as <š> in Hittite), Continental Celtic (but cf. 365–366), and Tocharian. In Balto-Slavic and Indo-Iranian, the so-called RUKI rule (p. 206, 495, 526–527) operates, according to which PIE \*s after \*r, \*u, \*k, \*i (and \*r, \*w, \*k<sup>w</sup>, \*y – in Ir. after \*k as well) turns to \*š, which then yields OInd. *ś* (written also as <š>), Av. *š* (p. 269), Lith. *š*, Latv./OPruss. *s*, Slav. *x* (cf. also PIE \*rs > *rš* in Arm. – p. 428). In Slavic, *x* (cf. CS *juxa*, Ved. *yúš-*, Lat. *iūs* – all ‘broth’, p. 33) can later be again palatalized to *š* (or *s*) (p. 528–529), while Lithuanian shows many exceptions from the law (p. 495). In Iranian, \*s generally changes to *h* (except in \*sC) – p. 264. In Greek (p. 293), \*sC and final \*-s# are preserved. Elsewhere, \*s yields *h*, written as <’> initially and preserved intervocalically in Mycenaean, but disappearing in Classical Greek (gen. sg. Gr. *νέφος* ‘cloud’, OCS *nebesē* ‘heaven’ < PIE \*neb<sup>h</sup>ese/os). Intervocalic \*s changes to *r* in Latin in historical times (p. 327); otherwise, it is unchanged. In Germanic, \*s becomes voiced word-finally (later again devoiced in Gothic) and medially in the case of Verner’s Law (p. 393). This \*z is preserved in East Germanic, but changes to *r* in North-West Germanic (cf. Goth. *batiza* to Eng. *better*). In Armenian (p. 427–428), PIE \*s is preserved in certain clusters (Arm. *mis*, OCS *měso* < PIE \*mēm̥s- ‘meat’) but, as in Greek, yields *h-* (or drops) initially (cf. Arm. *al* ‘salt’ ~ Lat. *sāl* but *hin* ‘old’ ~ Lat. *senex*) and disappears intervocalically. In Albanian, initially we get *gj-* (cf. Alb. *gjumë*, OCS *snb*, Gr. *ὑπνος* < PIE \*supnos ‘sleep’) or *sh-*, and medially perhaps *-sh-* or *-h-* (cf. p. 568).

### S-mobile

Some PIE roots appear both with initial \*s- and without it. This facultative \*s- is called *s-mobile* (mobile *s*). An example is PIE \*(s)poHimn- ‘foam’ > Ved. *phéna-*, OCS *pěny* (nom. pl.), OEng. *fām* (Eng. *foam*) but Lat. *spūma*, Lith. *spáinė*. *S-mobile* is usually explained through some kind of pre-PIE sandhi reanalysis (e.g., \*w<sub>l</sub>k<sup>w</sup>os pek<sub>y</sub>e(ti) ‘wolf sees’ > \*w<sub>l</sub>k<sup>w</sup>os (s)pek<sub>y</sub>e(ti) – cf. Lat. *speciō* ‘I see’ but Ved. *pásyati* ‘sees’).

### Consonantal reflexes of PIE laryngeals

The exact phonetics of PIE “laryngeals” (\*h<sub>1</sub>, \*h<sub>2</sub>, \*h<sub>3</sub>) are controversial. The name “laryngeals” is usually taken to be a traditional misnomer, though it is not impossible that \*h<sub>1</sub> was, for instance, indeed a laryngeal/glottal [h]. One possibility is that they were velar \*x̣, \*x, \*x<sup>w</sup> (parallel to stops), which would agree with their vowel-coloring characteristics (\*h<sub>2</sub>e > \*h<sub>2</sub>a and \*h<sub>3</sub>e > h<sub>3</sub>o, p. 43); the other that \*h<sub>1</sub> was glottal/laryngeal \*h or palatal \*ç,<sup>8</sup> that \*h<sub>2</sub> was a velar \*x (uvular \*χ or pharyngeal \*ħ are far less likely), and \*h<sub>3</sub> a labialized voiced velar \*ɣ<sup>w</sup> (or uvular \*ɣ<sup>w</sup>, for instance) or voiced pharyngeal \*ʕ (cf. the change \*ph<sub>3</sub> > \*b (p. 17) that would point to a voiced \*h<sub>3</sub>), etc. Other phonetic interpretations are also possible. The \*h<sub>2</sub> was the most frequent. The laryngeals disappear in most languages and positions, leaving only indirect traces (cf. p. 481, 523 for BSL), which is the reason they will be dealt with together with vowels (p. 43). Here, we shall only take a look at their consonantal reflexes (as *h* or aspiration).

In Anatolian, \*h<sub>2</sub> (and probably initial \*h<sub>3</sub>-) reflects as *h* in many positions. In Hittite, initial \*h<sub>2</sub>- yields *h*- (Hitt. *hartaggaš*, Gr. ἄρκτος < PIE \*h<sub>2</sub>r̥tkos ‘bear’). Medially, *-hh-* is preserved in many conditions (Hitt. *paḫḫur*, Gr. πῦρ < PIE \*peh<sub>2</sub>w<sub>r̥</sub> ‘fire’), but not all (p. 177). Many suppose that initial \*h<sub>3</sub>- also yields *h*- in Anatolian (Hitt. *ḫaštāi*, CLuw. *ḫāš-*, Ved. *ásthi*, Lat. *os* < PIE \*h<sub>3</sub>est<sub>h1</sub> ‘bone’), at least sometimes, though some contest that (and reconstruct PIE \*h<sub>2</sub>ost<sub>h1</sub> and the like instead). Medial \*-h<sub>3</sub>- generally disappears in Anatolian (but cf. p. 177).

In Indo-Iranian, laryngeals (mostly \*h<sub>2</sub>) aspirate preceding stops, both voiceless (p. 21) and voiced. Cf. PIE \*d<sup>h</sup>ugh<sub>2</sub>tēr ‘daughter’ > Ved. *duhitā* (with \*-gh<sub>2</sub>- > \*-g<sup>h</sup>- > *-h-*), Gr. θυγάτηρ (p. 47), \*megh<sub>2</sub> > Ved. *māhi* (p. 53, 80), or \*h<sub>1</sub>eḡHom > *ahām* (p. 82) (with \*ḡH > \*ḡ<sup>h</sup> > *h*). Initial \*h<sub>2/3</sub>- seems to yield *h*- in some cases in Armenian, though the exact conditions are unclear and not all scholars accept this (p. 430), and less convincingly in Albanian. Cf. Arm. *haw*, Lat. *avis* < PIE \*h<sub>2</sub>ew- ‘bird’ and Alb. *herdhe* ‘testicles’, Gr. ὄρχις ‘testicle’ < PIE \*h<sub>3</sub>erḡ<sup>h</sup>-.

The laryngeals apparently disappeared already in PIE medially before \*y (**Pinault’s Law**),<sup>9</sup> cf. PIE \*krew<sub>h2</sub>s (with *-i-* in Ved. *kravís-* and *-a-* in Gr. κρέας) (p. 16) but adjective \*krew<sub>h2</sub>yos > \*krewyos (Ved. *kravyás* ‘bloody’, cf. also the accent in Lith. *kraūjas*, which points to no laryngeal). For initial \*Hy- cf. p. 34. There were perhaps some other possible very specific PIE synchronic rules for laryngeal deletion – e.g., in \*VHTRV position (cf. PIE \*h<sub>2</sub>weh<sub>1</sub>- > Ved. *vāti* ‘blows’ but PGmc \*weðran ‘weather’, OCS *vedro* ‘nice weather’).

### PIE RESONANTS AND GLIDES

PIE segments were syllabic (vowels), asyllabic (stops, fricatives), or syllabic/asyllabic (resonants + glides). PIE asyllabic resonants were \*m, \*n, \*l, \*r and glides \*y, \*w. When in syllabic position, they are written as \*ṁ, \*ṅ, \*ḷ, \*ṛ and \*i, \*u (the last two were

functionally obviously vowels). Syllabic variants appear mostly in predictable positions: -CR̥C- (gen. sg. \*k̥r̥de/os ‘heart’), #R̥C- (gen. sg. \*udne/os ‘water’), and -CR̥# (\*yēk̥w̥r ‘liver’). In other positions, asyllabic variants appear: -VRC- (\*k̥ērd ‘heart’), -CRV- (\*ǵ<sup>h</sup>weh<sub>1</sub>r ‘beast’), #RV- (\*wod̥r), -VR# (\*ph<sub>2</sub>tēr ‘father’). However, the occurrence of an a-syllabic variant was not always completely predictable – cf. the initial resonant clusters like \*wr- (not \*\*ur-) or \*ml- (not \*\*ml-) (p. 52). The general rule in case of two potential syllabic resonants in a coda (-CRRC- or -CRR#) was that the second one became syllabic (“the \*R̥R̥ principle”), cf. instr. pl. \*k̥w̥ṇb<sup>h</sup>is ‘with dogs’ (Ved. *śvábhis*, not \*\*k̥unb<sup>h</sup>is) (p. 41). The exceptions here were cases with the present infix \*-n- (p. 95), which was always asyllabic (cf. \*yung- > Lat. *iungō* ‘I yoke’, \*link<sup>w</sup>- > Lat. *linquō* ‘I leave’ and not \*\*iwn̥g-, \*\*lyṇk<sup>w</sup>-); accusatives \*-im/-um (not \*\*-ym̥/-wṇm̥, p. 72–73) and pl. \*-ins/-uns/-tr̥ns (not \*\*-yṇs/-wṇs/-tr̥s, p. 72, 74–75); and cases with Stang’s Law like the acc. sg. \*dyewm (not \*\*dyewṇ) > \*dyēm (p. 72). All this means that syllabic resonants were really separate phonemes in the last phase of PIE and not just allophones of their asyllabic variants (although still regularly alternating with them in many cases). In any case, the reflexes of syllabic and asyllabic resonants are completely divergent and have to be analyzed separately.

### Asyllabic resonants

Asyllabic resonants were mostly stable in IE languages. They remained unchanged in most languages and positions.

#### \*m

**PIE \*men-** ‘think’ > Ved. *mánas-* ‘mind’ (YAv. *manah-* ‘thinking’), Gr. μένος ‘spirit’, Lat. *meminī* ‘I remember’ (with root reduplication, ⇒ Eng. *memory*), OCS *mněti* ‘to think’, Lith. *menù* ‘I remember’, Goth. *man* ‘to think’, OIr. *muinithir* ‘thinks’, Arm. *imanam* ‘I understand’, Toch. B *mañu* ‘desire’

**PIE \*med<sup>h</sup>u** ‘mead’ > Ved. *mádhu* (YAv. *maðu-* ‘berry wine’), Gr. μέθυ ‘wine’, OCS *medь* ‘honey’, Lith. *medùs* ‘honey’, OEng. *me(o)du* (Eng. *mead*), OIr. *mid* (Gaul. *Medu-genos*, Clb. *Mezu-kenos* (pers. names)), Toch. B *mīt* ‘honey’

**PIE \*med<sup>h</sup>y-** ‘middle’ > Ved. *mádhyas* (GAv. *ma’dya-*), Gr. μέσ(σ)ος (⇒ Eng. *Mesozoic*), Lat. *medius* (⇒ Eng. *medium*), OCS loc. pl. *meždaxь* ‘alleys’ (PSl. \*medja ‘boundary’), Lith. *mėdė* ‘forest’ (< \*‘forest as a border’), Goth. *midjis* (Eng. *mid(dle)*), OIr. *mide*, Arm. *měj*, Alb. *midis* ‘in the middle of’

Generally speaking, PIE \*m does not change in any of the languages. The typologically frequent (cf. Arabic, Finnish) change of final -m# > -n# occurs in many languages, e.g., in Greek, Hittite, Germanic, Armenian, Old Prussian, Phrygian, and many Celtic languages (cf. PIE neuter nom. sg. \*-om > OInd. -am, Lat. -um but Gr. -ov, Hitt. -an, PGmc \*-an, OPruss. -an, p. 65).

#### \*n

**PIE \*neb<sup>h</sup>os** ‘cloud, cloudy sky, mist’ > Ved. *nábhās* ‘cloud, mist’ (GAv. *nabah-* ‘cloud’), Gr. νέφος ‘cloud’, Lat. *nebula* ‘mist, fog’ (⇒ Eng. *nebula*), OCS *nebo* ‘heaven’ (Slav. also ‘sky’), Lith. *debesis* ‘cloud’ < \*nebesis, OHG *nebul* ‘fog, mist’, Hitt. *nēpiš* ‘sky, heaven’, OIr. *nem* ‘heaven, sky’

**PIE \*new(y)os** ‘new’ > Ved. *náv(y)as* (YAv. *nauua-*), Gr. *véος* (⇒ Eng. *neo-*), Lat. *nouus*, OCS *novъ*, Lith. *naĩjas*, Goth. *niujis* (Eng. *new*), Hitt. *nēwa-*, OIr. *núae* (Gaul. *Nouio-dunum* (toponym)), Arm. *nor*, Toch. B *ñuwe*

**PIE \*nog<sup>h</sup>ts** ‘night’ (p. 54) > Ved. *nákt-* (YAv. *upa-naxtar-* ‘almost night’), Gr. *νύξ*, Lat. *nox* (⇒ Eng. *nocturnal*), OCS *noštъ*, Lith. *naktis*, Goth. *nahts* (Eng. *night*), Hitt. *nekuz mēhur* ‘at night’, OIr. *innocht* ‘tonight’ (Gaul. *tri-nox(ion)* ‘three nights’), Alb. *natë*, Toch. A *nokte* ‘at night’

Generally speaking, PIE \*n does not change in any of the languages, but cf. late \*VnV > r in Tosk Albanian (p. 564–565) and Slavic nasal vowels (p. 521). In cases like \*peŋk<sup>w</sup>e (p. 24), PIE probably had an allophone \*[ŋ] before velars like most languages.

### \*r

**PIE \*swesōr** ‘sister’ > Ved. *svásar-* (YAv. *x<sup>w</sup>an̥har-*), Gr. (Hesychius) *ἔορ* ‘daughter, first cousin’, Lat. *soror* (⇒ Eng. *sorority*), OCS *sestra*, Lith. *sesuo* – gen. sg. *sesešs*, Goth. *swistar* (Eng. *sister*), OIr. *siur* (Gaul. instr. pl. *suiorebe*), Arm. *k<sup>o</sup>oyr*, Toch. B *šer*

**PIE \*treyes** ‘three’ > Ved. *tráyas* (YAv. *θrāyō*), Gr. *τρεῖς* (⇒ Eng. *tri-logy*), Lat. *trēs*, OCS *trije*, Lith. *tryš*, Goth. *þreis* (Eng. *three*), Hitt. *teri-*, OIr. *trí* (Gaul. *tri-*), Arm. *erek<sup>ε</sup>*, Alb. *tre*, Toch. B *traí*

**PIE \*mer-** ‘disappear > die’ > Ved. *maranti* ‘they die’ (YAv. *fra-miryete* ‘dies’), Gr. *μορτός* ‘mortal’, Lat. *moriōr* ‘I die’ (⇒ Eng. *mortal*), OCS *mrěti* ‘to die’, Lith. *mir̃ti* ‘to die’, Goth. *maúrþr* (Eng. *murder*), Hitt. *merzi* ‘disappears’, Arm. *meřanim* ‘I die’

Since it had a syllabic variant (\*r̥ – p. 35), PIE \*r was originally phonetically probably a trill (as it still is, e.g., in Slavic). Generally speaking, PIE \*r does not change in any of the languages – except phonetically to a different type of rhotics (rarely, one finds an *l* as a reflex in some Indo-Iranian languages, p. 206, 271). It is possible that there was no initial \*#r- in PIE (as in Proto-Uralic, most non-IE Caucasian languages, Korean, etc.), which would be preserved as a phonotactic constraint in Old Anatolian languages, Armenian, Phrygian, and Greek (for Gr. #r̥- [hr̥] < \*wr-, \*sr-, cf. p. 290). If so, an initial laryngeal would have to be reconstructed even when there is no direct attestation of it, e.g., PIE \*HrotHos ‘wheel’ (Ved. *ráthas* ‘chariot’, Lat. *rota*, Lith. *rātas*).

### \*l

**PIE \*leyǵ<sup>h</sup>-** ‘lick’ > Ved. *rédhi* (younger), *leđhi* ‘licks’ (YAv. *riz-*, Modern Pers. *lištan* ‘to lick’), Gr. *λείγω* ‘I lick up’, Lat. *lingō* ‘I lick’, OCS *lizaaše* ‘licked’, Lith. *ližėti* ‘to lick’, OEng. *liccian* ‘to lick’ (Eng. *lick*), OIr. *ligid* ‘licks’, Arm. *lizem* ‘I lick’

**PIE \*seh<sub>2</sub>wl-** ‘sun’ (cf. p. 78) > Ved. *sūryas* (GAv. *huuarō*), Gr. *ἥλιος* (⇒ Eng. *helium*), Lat. *sōl* (⇒ Eng. *solar*), OCS *slěnyce*, Lith. *saulė*, Goth. *sauil*, MW *haul*

**PIE \*h<sub>3</sub>el(n)-** ‘elbow’ > Ved. *aratnís* (YAv. *frārāñni-drājah-* ‘elbow (length)’), Gr. *ὠλένη*, Lat. *ulna*, OCS *lakъъ*, Lith. *alkūnė*, Goth. *aleina* ‘cubit’ (Eng. *ell*, *el-bow*), OIr. *uilen* ‘elbow, angle’ (Gaul. *Olina* (river name)), Arm. *olok<sup>ε</sup>* ‘shin’, Alb. (Gheg) *llānë* ‘fore-arm’, Toch. B *alyiye* ‘palm of the hand’

Generally speaking, PIE \*l is preserved in most languages. The only exception is Indo-Iranian, where it yields *r* in most languages (and in most words), but not always (p. 206, 271).

## \*w

**PIE \*weg<sup>h</sup>-** ‘drive, convey’ > Ved. *váhati* ‘drives’ (YAv. *vazənti* ‘they pull’), Gr. ὄχος ‘carriage’ (Myc. *wo-ka* ‘wagon’, Pamph. *φεχέτω* ‘should bring’, Cypr. *ἔφεξε* ‘brought’), Lat. *uehō* ‘I bear, carry, convey’ (⇒ Eng. *vector*, *vehement*), OCS *vezō* ‘I drive, lead, convey’, Lith. *vėžti* ‘to lead, convey’, Goth. *ga-wigan* ‘to move’ (Eng. *weigh*; Eng. *wagon* ← Dutch), OIr. *fén* ‘wagon’ (W *gwain*), (?) Alb. *vjedh* ‘rob, steal’, Toch. A *wkāṃ* ‘way, manner’

**PIE \*wert-** ‘turn, spin’ > Ved. *vártate* ‘turns’ (YAv. *var<sup>o</sup>ta-* ‘turn’), (?) Gr. (Dor., Hesychius) *πάταναν* *τορύνην* ‘stirrer, ladle’, Lat. *uertō* ‘I turn’ (⇒ Eng. *con-vert*), OCS *vratiti se* ‘to return, turn’, Lith. *veĩsti* ‘to turn over’, Goth. *wairþan* ‘to become’, OIr. *frith-* ‘against’ (cf. Lat. *uersus* ‘against, towards’ ⇒ Eng. *ad-verse*), Toch. A *wärt-* ‘throw’

**PIE \*welh<sub>1</sub>-** ‘want’ > Ved. *várat* ‘wants, chooses’ (GAv. *və<sup>o</sup>ntē* ‘chooses’), Gr. *λῶ* ‘I want/wish’, Lat. *uolō* ‘I wish/want’ (⇒ Eng. *volunteer*), OCS *velěti* ‘to want’, OLith. *vėlti* ‘to let, allow’, Goth. *wiljan* ‘to want’ (Eng. *will*), Gaul. *uelor* ‘I wish’ (MW *gwell* ‘better’)

Generally speaking, PIE \*w is preserved in many languages in most positions. However, the original PIE bilabial semivowel pronunciation (like Eng. *w*) is preserved only in some languages (Classical Latin but not Romance, some Ancient Greek dialects, English but not most Germanic languages, etc.), while elsewhere it usually changes to a fricative *v* (Lithuanian, most Slavic languages, Albanian, most Germanic and Romance languages, etc.). A labiodental approximant [ʋ] as a reflex is less common (as in some Dutch and South Slavic Štokavian dialects, Slovene, as one of the realizations in Hindi, etc.).

Mycenaean still has a *w*, and so do some Greek dialects (written as *Ϝ*); this *Ϝ* disappears early in Attic-Ionic (p. 289, 292). PCelt. \*w is preserved in Gaulish and Celtiberian (later changed in Insular Celtic, p. 362, 365–366). In Armenian, -w is preserved word-finally (cf. Arm. *naw* ‘ship’, Lat. *nāuis* < PIE \*neh<sub>2</sub>w-), while initially it changes to *g* (via \*gw, similar to Welsh), cf. Arm. *egit* ‘found’, Ved. *ávidat* ‘found’ < PIE \*widet (p. 429). In Albanian, PIE \*w is preserved initially, and otherwise it disappears (p. 568).

## \*y

**PIE \*yuh<sub>3</sub>s-** ‘soup, broth’ > Ved. *yús*, Gr. ζύμη ‘leaven’, ζωμός ‘soup’, Lat. *iūs* ‘broth, sauce’, CSL *juxa*, Lith. arch. *jūšė* ‘fish soup’

**PIE \*h<sub>2</sub>yowHn-** ‘young’ > Ved. *yúvan-* (YAv. *yauua* ‘youth’), Lat. *iuuenis* ‘young man’ (⇒ Eng. *juvenile*, *junior*), OCS *junь*, Lith. *jáunas*, Goth. *juggs* (Eng. *young*), OIr. *óac* (MW *ieuanc*, Gaul. *Iouinc-illus* (pers. name))

**PIE \*yēk<sup>h</sup>r̥** ‘liver’ (p. 78) > Ved. *yákr̥t* (YAv. *yākarə*), Gr. ἥπαρ (⇒ Eng. *hepatitis*), Lat. *iecur*, CS *ikra* ‘roe’ (< PSI. \*jьkra), OLith. *jėknoš*

Generally speaking, PIE \*y is often preserved but not universally. In Greek and Latin, intervocalic \*y disappears (PIE \*treyes ‘three’ > Gr. τρεῖς [trēs], Lat. *trēs*), \*-Cy- has

various outcomes (p. 293), and in initial position one finds either Gr.  $\epsilon$ -/h-/ or  $\zeta$ -, which is an old and disputed problem – some claim that one is the outcome of  $*y$ - and the other of  $*Hy$ - (since there are no “prothetic vowels” – p. 48 – before  $*y$ ), but which reflex is derived from which source is also disputed. PCelt.  $*y$  disappears completely in Old Irish (p. 362, 365–366). In Armenian, intervocalic  $*y$  disappears (PIE  $*treys$  > Arm. *erek<sup>c</sup>*),  $*-Ry$ - yields *Rj* (cf. Arm. *olj*, OIr. *uile* < PIE  $*h_3elyos$  ‘whole’), and the reflex of initial  $*y$ - is not completely clear (p. 429). In Albanian, medial  $*y$  is lost (PIE  $*treys$  > Alb. *tre*), while initial  $*y$ - yields *gj*-, cf. Alb. *gjes* ‘gird’, Gr.  $\zeta\omega\sigma\tau\acute{\eta}\rho$  ‘a warrior’s belt’, Lith. *júosta* ‘waist-band, belt’ < PIE  $*yeh_3s$ - (p. 568).

TABLE 1.4 THE REFLEXES OF PIE ASYLLABIC RESONANTS

PIE	OInd.	Av.	Gr.	Lat.	OCS	Lith.	Goth.	Hitt.	OIr.	Arm.	Alb.	Toch.
$*m$	<i>m</i>	<i>m</i>	$\mu$ , -v#	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i> , -n#	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>
$*n$	<i>n</i>	<i>n</i>	<i>v</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
$*l$	<i>r/l</i>	<i>r</i>	$\lambda$	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>
$*r$	<i>r</i>	<i>r</i>	$\rho$	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r(r)</i>	<i>r</i>
$*w$	<i>v</i>	<i>v</i>	$\text{f} > \#0\text{-/}\text{'}\text{-}$ , VØV	<i>u</i>	<i>v</i>	<i>v</i>	<i>w</i>	<i>w</i>	$\#f$ , VØV, Cb	<i>g</i> , -w#	$\#v\text{-}$ , Ø	<i>w</i>
$*y$	<i>y</i>	<i>y</i>	$\#\text{-}/\zeta\text{-}$ , VØV	<i>i</i> , VØV	<i>j</i>	<i>j</i>	<i>j</i>	<i>y</i>	Ø	-Rj-, VØV	$\#gj\text{-}$ , -Ø-	<i>y</i>

### Syllabic resonants

Syllabic resonants  $*m$ ,  $*n$ ,  $*l$ ,  $*r$  are completely unstable diachronically in IE and change everywhere (the sole exception being PIE  $*r$  > OInd. *r*). Their reflexes are of the following types:

- the reflex is a syllabic resonant ( $*R > R$ ) – only in PIE  $*r$ ,  $*l$  > OInd. *r*
- the resonant obtains a preceding vowel ( $*R > VR$ ) – e.g., PIE  $*R > \text{Gmc } *uR$
- the resonant obtains a following vowel ( $*R > RV$ ) – e.g., PIE  $*r$  > Gr.  $\rho\alpha$  (also  $\alpha\rho$ )
- the resonant becomes a vowel ( $*R > V$ ) – e.g., PIE  $*m$ ,  $*n$  > OInd./Alb. *a*, Gr. *a*

The type b) is the most frequent by far. In some languages/groups, all resonants have the same type of development (cf. Gmc  $*um/un/ul/ur$ ). In others, the type of development of  $*m/n$  differs from that of  $*l/r$  (cf. Lat. *em/en* but *ol/or*). Reflexes in  $*-CR_{\text{H}}V$ - (before a disappearing laryngeal) are often different than in  $*-CR_{\text{C}}(V)$ - (p. 39). As already said,  $*i$  and  $*u$  phonologically acted just like syllabic resonants.

### $*m$

**PIE  $*dek_{\text{m}}$**  ‘ten’ > Ved. *dása* (YAv. *dasa*), Gr.  $\delta\acute{\epsilon}\kappa\alpha$  ( $\Rightarrow$  Eng. *decade*), Lat. *decem* ( $\Rightarrow$  Eng. *December*), OCS *desętb*, Lith. *dęšimt*, Goth. *taihun* (Eng. *ten*), OIr. *deich<sup>n</sup>* (Gaul. *decan-*), Arm. *tasn* <  $*tasan$

**PIE  $*(d)k_{\text{m}}t_{\text{om}}$**  ‘hundred’ > Ved. *śatám* (YAv. *satám*), Gr.  $\acute{\epsilon}\kappa\alpha\tau\acute{\omicron}\nu$ , Lat. *centum* ( $\Rightarrow$  Eng. *cent*), OCS *stęto*, Lith. *šimtąs*, Goth. *hunda* (Eng. *hundred*), OW *cant* (Cib. *kantom*, OIr. *cét*), Toch. A *kánt*

**PIE  $*g^{w}m_{\text{tis}}$**  ‘walk’ > Ved. *gátis* (YAv. *a<sup>i</sup>βi-ga<sup>i</sup>ti* ‘entrance’), Gr.  $\beta\acute{\alpha}\sigma\iota\varsigma$  ‘step’ ( $\Rightarrow$  Eng. *basis*), Lat. *uentiō* ‘coming’ ( $\Rightarrow$  Eng. *con-vention*), Goth. *ga-qumþs* ‘gathering’

**\*ŋ**

**PIE \*mŋtis** ‘thought, mind’ > Ved. *matís* ‘thought’ (Av. *ma’ti-*), Gr. ἀυτό-ματος ‘spontaneous’ (⇒ Eng. *automatic*), Lat. (*mēns*) – gen. sg. *mentis* ‘mind’ (⇒ Eng. *mental*), OCS *pa-mętb* ‘memory’ (Slav. also ‘mind’), Lith. *mintis* ‘thought’, Goth. *ga-munds* ‘remembrance’ (Eng. *mind*)

**PIE \*h<sub>1</sub>newŋ** ‘nine’ > Ved. *náva* (YAv. *nava*), Gr. ἐννέα, Lat. *nouem* < \*nouen (by analogy to *decem*, ⇒ Eng. *november*), OCS *devętb*, Lith. *deviņtas* ‘ninth’ (OPruss. *newīnts*), Goth. *niun* (Eng. *nine*), OIr. *noí<sup>n</sup>* < PCelt. \*nowan, Arm. *inn* < \*inan

**PIE \*h<sub>3</sub>neh<sub>3</sub>mŋ** ‘name’ (p. 48) > Ved. *náma* (YAv. *nama*), Gr. ὄνομα (⇒ Eng. *onomastics*), Lat. *nōmen* (⇒ Eng. *nominal*), Hitt. *lāman*

The reflexes of \*ŋ and \*ŋ̥ are always parallel. Final *-m* changes to *-n* in some languages as usual (p. 31), and assimilatory *m* > *n* before dentals is also common in some languages. In Indo-Iranian, Albanian, and Greek, the reflex is *a/α* (cf. Alb. *shta-të* ‘seven’ < \*s(ep)ta- < PIE \*septŋ). Latin has *em/en* (initially *ma-*, *na-*); Germanic \*um/un; Armenian, Hittite, and Celtic *am/an* (with further development in separate Celtic languages – p. 358–359); and Tocharian *ām/ān*. In Balto-Slavic, the reflexes are \*im/in (less frequently and in unclear conditions \*um/un). Lithuanian preserves *im/um*, *in/un*, and in Slavic \*im/in > *e*, \*um/un > *o* (\*um can less often and in unclear conditions also yield *u*, cf. ‘hundred’ above) in closed syllables.

**\*l̥**

**PIE \*wl̥k<sup>w</sup>os** ‘wolf’ > Ved. *v̥́kas* (YAv. *vəhrka-*), OCS *vl̥bkъ* < PSI. \*v̥lkъ, Lith. *vilkas*, Goth. *wulfs* (Eng. *wolf*), (?) Hitt. *walkuwa-* (meaning something negative), Toch. B *walkwe*

**PIE \*ml̥dus** ‘soft’ > Ved. *ḿ̥dús*, Gr. βλαδέϊς ‘powerless’, Lat. *mollis* < \*moldwis, MW *blydd* < PCelt. \*mlidos

**PIE \*pl̥th<sup>w</sup>us** ‘broad, wide’ > Ved. *pý̥thús* (GAv. *pəṛ<sup>w</sup>θu-*), Gr. πλατύς, OBret. *litan* ‘broad’ (Gaul. *Litana* (toponym)), Arm. *y-alt<sup>w</sup>*

**\*r̥**

**PIE \*mrt(w)os** ‘dead’, \*ḿ̥tis ‘death’ (cf. \*mer-, p. 32) > Ved. *ḿ̥tás* ‘dead’ (YAv. *məṛ<sup>w</sup>ta-*), Gr. (Aeol.) βροτός ‘mortal’, Lat. *mors* ‘death’ (⇒ Eng. *mortal*), *mortuus* ‘dead’, OCS *mr̥tvъ* ‘dead’ < \*ḿ̥rtvъ, *sъ-mr̥tvъ* ‘death’ < \*sъ-ḿ̥rtvъ, Lith. *mirtis* ‘death’, Goth. *maúrþr* (Eng. *murder*), Arm. *mard* ‘man’ < \*‘mortal’, MW *marw* ‘dead’

**PIE \*h<sub>2</sub>rt̥kos** > \*h<sub>2</sub>rt̥kos [-k̥b] ‘bear’ (p. 53) > Ved. *ṛ̥́k̥sas* (YAv. *arša-*), Gr. ἄρκτος (⇒ Eng. *Arctic*), Lat. *ursus*, Lith. *irštva* ‘bear-den’, OIr. *art* (Gaul. *Artio* (name of a deity)), Hitt. *ḫartaggaš*, Arm. *arj*, Alb. *ari*

**PIE \*k<sup>w</sup>rmis** ‘worm’ > Ved. *ḱ̥mis*, OCS *čr̥tvъ* < \*č̥r̥tvъ, Lith. *kirmis*, OIr. *cruim* < PCelt. \*k<sup>w</sup>rimis, Alb. *krim(b)* (cf. PIE \*w̥rmis > Goth. *waúrm* ‘snake’ (Eng. *worm*))

The reflexes of \*l̥ and \*r̥ are always parallel. In Old Indic we find *ṛ*, in Avestan mostly *əṛ<sup>w</sup>* (p. 267). In Greek (Attic), we find *αλ/λα* and *αρ/ρα* in unclear conditions (cf. PIE \*k̥rd-, p. 21); in Latin *or/ol* (initially *ra-*, *la-*). Balto-Slavic has \*ir/il (less frequently



\*ur/ul) – unchanged in Lithuanian; Slavic reflexes are \*ьr/ьr and \*ьl/ьl (p. 480), preserved in Old Russian but written as <ръ/ръ> and <ль/ль> indiscriminately in OCS. Germanic has \*ul/ur (> *ull/aúr* in Gothic, p. 391). In Celtic we find \*ri/li before stops and \*m, and \*ar/al otherwise (p. 362). Armenian and Hittite have *ar*, *al*, Albanian usually *ri* and *li*, and Tocharian *äl*, *är*.

## \*i

**PIE \*misd<sup>h</sup>-** ‘reward, pay’ > Ved. *mīdhám* < \*mizdhám ‘spoils, gain’ (MPers. *mizd*), Gr. μισθός ‘hire, pay, reward’, OCS *mъzda* ‘pay’, Goth. *mizdo*

**PIE \*h<sub>1</sub>esmi** ‘am’ > Ved. *ásmi* (YAv. *ahmi*), Gr. εἰμί, OCS *jesmь*, Hitt. *ēšmi*

**PIE \*-is** (nom. sg., *i*-stems, p. 72) > Ved. *-is*, Gr. *-ις*, Lat. *-is*, OCS *-ь*, Lith. *-is*, Hitt. *-iš*

## \*u

**PIE \*yugom** ‘yoke’ > Ved. *yugám* (GAv. *yaogaṭ* ‘harnesses’), Gr. ζυγόν, Lat. *iugum* (⇒ Eng. *conjugation*), OCS *igo* < \*jъgo, Lith. *jūngas*, Goth. *juk* (Eng. *yoke*), Mlr. *cuing* < \*kom-yungi- (Gaul. *Ver-iugus* (pers. name)), Hitt. *iukan*, Arm. *luc*

**PIE \*snusos** ‘daughter-in-law’ > Ved. *snuṣá*, Gr. (Hom.) νύος, Lat. *nurus*, CS *snъxa*, OHG *snura*, Arm. *nu*

**PIE \*supnos** ‘sleep’ > Pali *supina-* ‘dream’, Gr. ὕπνος (⇒ Eng. *hypnosis*), OCS *сънъ*, Alb. *gjumë*

In most languages/families, PIE \*i and \*u are preserved (at least in the oldest stages). Greek υ originally had the value [u] but changed to [ü] later in Attic (p. 295). In Slavic, \*u > ъ and \*i > ѝ (originally probably back and front very short, centralized/schwa-like vowels respectively). In Gothic, \*u changes to *au* [ɔ] (written originally just as <au>, the same as the diphthong [aw]) before *h*, *hv*, *r*. In Tocharian, both \*i (also \*e) and \*u yield *ä*, but *ä* from \*i/e causes palatalization (p. 455, 459).

TABLE 1.5 THE REFLEXES OF PIE SYLLABIC RESONANTS

PIE	OInd.	Av.	Gr.	Lat.	OCS	Lith.	Goth.	Hitt.	PCelt.	Arm.	Alb.	Toch.
*ṛ̥	<i>a</i>	<i>a</i>	<i>α</i>	<i>em, #ma-</i>	<i>ę, ɔ</i>	<i>iṛ̥/uṛ̥</i>	<i>um</i>	<i>am</i>	*am	<i>am</i>	<i>a</i>	<i>äm</i>
*ṝ̥	<i>a</i>	<i>a</i>	<i>α</i>	<i>en, #na-</i>	<i>ę, ɔ, ъ</i>	<i>iṝ̥/uṝ̥</i>	<i>un</i>	<i>an</i>	*an	<i>an</i>	<i>a</i>	<i>än</i>
*ṝ̥̄	<i>r̥</i>	<i>əṛ̥</i>	<i>αλ/λα</i>	<i>ol, #la-</i>	<i>ь/ь̄</i>	<i>iṝ̥̄/uṝ̥̄</i>	<i>ul</i>	<i>al</i>	*al, *li/-C <sub>stops</sub> /m	<i>al</i>	<i>li</i>	<i>äl</i>
*ṝ̥̄̄	<i>r̥</i>	<i>əṛ̥</i>	<i>αρ/ρα</i>	<i>or, #ra-</i>	<i>рь/рь̄</i>	<i>iṝ̥̄̄/uṝ̥̄̄</i>	<i>aúr</i>	<i>ar</i>	*ar, *ri/-C <sub>stops</sub> /m	<i>ar</i>	<i>ri</i>	<i>är</i>
*i	<i>i</i>	<i>i</i>	<i>ι</i>	<i>i</i>	<i>ь</i>	<i>i</i>	<i>i</i>	<i>i</i>	*i	<i>i</i>	<i>i</i>	<i>'ä</i>
*u	<i>u</i>	<i>u</i>	<i>υ</i>	<i>u</i>	<i>ъ</i>	<i>u</i>	<i>u</i>	<i>u</i>	*u	<i>u</i>	<i>u</i>	<i>ä</i>

## Syllabic resonants and laryngeals

In traditional IE linguistics, PIE reconstruction included long syllabic resonants \*ṝ̥̄̄, \*ṝ̥̄̄̄, \*ṝ̥̄̄̄̄, \*ṝ̥̄̄̄̄̄, \*ī, and \*ū (in \*CR̥C position). After the discovery of laryngeals, it turned out that these were in fact \*ṝ̥̄̄̄H, \*ṝ̥̄̄̄̄H, \*ṝ̥̄̄̄̄̄H, \*īH, \*ī̄H, and \*ūH. One has to distinguish two basic positions of \*ṜH – one is the pre-consonantal (\*CR̥̄HC) and the other is the pre-vocalic

(\*C<sub>ṛ</sub>HV). In \*C<sub>ṛ</sub>HC, what one gets in IE languages (if the reflexes are not the same as those of the regular syllabic resonants) are sequences like R<sub>V</sub>,  $\bar{V}$ R, or, less frequently,  $\bar{V}$ R<sub>V</sub>. In \*C<sub>ṛ</sub>HV, the reflexes are everywhere of the  $\bar{V}$ R type. We deal first with the \*C<sub>ṛ</sub>HC type.

\*C<sub>ṛ</sub>HC (pre-consonantal position)

\***ṛH**

PIE \***dmh<sub>1</sub>tos** ‘tamed’ > Ved. *dāntás* (with analogical *-n-*), Gr. ἀδάμα(σ)τος ‘unconquered’, ἄδμητος ‘unbroken’

\***ṛH**

PIE \***ǵṛh<sub>1</sub>tos** ‘born’ > Ved. *jātás* (YAv. *zāta-*), Gr. κασί-γνητος ‘brother’, Lat. (*g*) *nātus*, Goth. *airþa-kunds* ‘born of the earth’, Gaul. *Cintu-gnātus* (‘first-born’ (pers. name))

PIE \***ǵṛh<sub>3</sub>tos** ‘known’ > Gr. γνωτός, Lat. *gnārus* ‘knowing’ (< \*ǵṛh<sub>3</sub>ros), Lith. *pa-žin-tas*, Goth. *-kunþs*, OIr. *gnáth*

There are few words with \*ṛH/ṛH (mostly \*-tos participles of verbs of \*CVRH- type), especially with \*ṛH. In Vedic, *ā* is the reflex (though this is not completely certain for \*ṛH). In Greek, the reflexes depend on the laryngeal and possibly (though disputably) on stress position. In unaccented position, \*ṛh<sub>1</sub> > μη, \*ṛh<sub>2</sub> > μᾶ (> Att. μη, p. 295), \*ṛh<sub>3</sub> > μω and \*ṛh<sub>1</sub> > νη, \*ṛh<sub>2</sub> > νᾶ (> Att. νη), \*ṛh<sub>3</sub> > νω. When stressed, the reflexes seem to be ἐμε/άμα/όμο and ἐνε/άνα/όνο. Cf. pairs like πολύκιμητος ‘wrought with much toil’ – κάματατος ‘toil’ (\*k<sub>ṛ</sub>h<sub>2</sub>tos) or κασίγνητος and γένεσις ‘origin’ (\*ǵṛh<sub>1</sub>tis). But forms like γνήσιος ‘belonging to the race’ would then have to be secondary. In Germanic, the reflexes are identical to those of \*ṛH/ṛ, and the same goes for Balto-Slavic (disregarding the prosodical differences, see below). Latin has *mā*, *nā*, and so does Celtic (OIr. *má*, *ná*; cf. here also Toch. <*mā*>, <*nā*>), while in Armenian we find *ama*, *ana* (somewhat similar to Greek).

\***ṛH**

PIE \***plh<sub>1</sub>nos** ‘full’ > Ved. *pūrṇás* (GAv. *pəṛna-*), OCS *plъnъ* (BCMS *pŭn*), Lith. *pilnas*, Goth. *fulls* (Eng. *full*), OIr. *lán*

PIE \*(**h<sub>2</sub>**)**wl̥h<sub>1</sub>neh<sub>2</sub>** ‘wool’ > Ved. *úrṇā* (YAv. *varnā-*), Gr. λῆνος, Lat. *lāna*, OCS *vlъna* (BCMS *vŭna*), Lith. *vilna*, Goth. *wulla* (Eng. *wool*), MW *gwlan*

PIE \***d̥l̥h<sub>1</sub>g<sup>h</sup>os** ‘long’ > Ved. *dīrghás* (GAv. *dar̥ga-*), OCS *dlъgъ* (BCMS *dŭg*), Lith. *ilgas* (< \*dilgas), Goth. *tulgus* ‘firm’, Alb. *gjatë* (< \*glag(V)-t)

\***ṛH**

PIE \***ǵṛh<sub>2</sub>no-** ‘old, ripe; grain’ (cf. \*ǵerh<sub>2</sub>-, p. 22) > Ved. *jūrṇás* ‘old’ (YAv. *zar̥ta-* ‘decrepit’), Lat. *grānum* ‘grain’ (⇒ Eng. *grain*), OCS *zrъno* ‘grain’ (BCMS *zŕno*), Lith. *žirnis* ‘grain’, Goth. *kaurn* ‘grain’ (Eng. *corn*), OIr. *grán* ‘grain’

- PIE \*k<sub>ṛ</sub>h<sub>2</sub>srō(n)** ‘hornet’ > Lat. *crābrō*, CS *srъšenъ*, Lith. *širšė*, OEng. *hyrnetu* (Eng. *hornet*) < PGmc \*xurzn-  
**PIE \*str<sub>h</sub>3t/nos** ‘laid down’ > Ved. *sīrṇás*, Gr. στρωτός, Lat. *strātus* (⇒ Eng. *street*)

PIE \*ṛH/ḲH merge and yield *ūr/ūr* in Vedic (*ūr* usually after labials – *p, ph, bh, m, v*) and *ar<sup>o</sup>* in Avestan. In Greek once again we find λη/ρη, λᾱ (> Att. λη)/ρᾱ, λω/ρω and ἐλε/ερε, ἄλα/ἄρα, ὄλο/όρο. In Germanic \*ṛ/Ḳ = \*ṛH/ḲH; the same is true for Balto-Slavic except on prosodic level – cf. BCMS *pūn*, Lith. *pilnas*, Latv. *piļns* for PIE \*p<sub>ḷ</sub>h<sub>1</sub>nos (p. 37), but BCMS *vūk*, Lith. *vilkas*, Latv. *vilks* for PIE \*w<sub>ḷ</sub>k<sup>w</sup>os (p. 35). Latin and Celtic have *lā, rā* (written as *lá, rá* in OIr. – p. 362 – cf. also Toch. <lā>, <rā>), while in Armenian we find *ala, ara*. In Albanian we find possible *al, ar* reflexes (cf. Alb. *parë* ‘first’ < PIE \*p<sub>ṛ</sub>h<sub>2/3</sub>wo-) but also *la, ra* (cf. *gjatë* above).

### \*iH

- PIE \*wiHros** ‘man’ > Ved. *vīrás* (Av. *vīra-*), Lat. *uir* (with shortening, cf. p. 357), Lith. *výras*, Goth. *wair* (with shortening, cf. Eng. *were-wolf*), OIr. *fer* (with shortening), Toch. A *wir* ‘young’  
**PIE \*g<sup>w</sup>ih<sub>3</sub>wos** ‘alive’ > Ved. *jīvás*, Gr. βίος ‘life’ (with pre-vocalic shortening, ⇒ Eng. *bio-logy*, etc.), Lat. *uīuus* (⇒ Eng. *vivi-section*), OCS *živъ*, Lith. *gývas*, Goth. *qius* ‘alive’ (with shortening, Eng. *quick*), OIr. *béo* (Ogham BIVI-TI, with shortening)  
**PIE \*g<sup>w</sup>riHweh<sub>2</sub>** ‘neck’ > Ved. *grīvā* (YAv. *grīvā*), OCS *grivъna* ‘necklace’ (BCMS *grīva* ‘mane’), Latv. *grīva* ‘(river) mouth’

### \*uH

- PIE \*suHnus** ‘son’ > Ved. *sūnús* (GAv. *hunu-* ‘offspring’), Gr. dial. υἱός (< \*suHyus), OCS *synъ*, Lith. *sūnūs* (acc. sg. *sūnu*), Goth. *sunus* (with shortening, Eng. *son*)  
**PIE \*muHs** ‘mouse’ > Ved. *mūṣ-*, Gr. μῦς, Lat. *mūs*, OCS *myšъ* (BCMS *miš*), OEng. *mūs* (Eng. *mouse*), Arm. *mukn*, Alb. *mi*  
**PIE \*suHs** ‘sow, swine’ > Ved. *sūkarás* ‘wild boar’ (YAv. *hū-* ‘swine’), Gr. ὄς, Lat. *sūs* ‘pig, sow’, OEng. *sú* ‘sow’, Alb. *thi* ‘pig’ (with secondary *th-*), Toch. B *suwo* ‘pig, hog’

Traditionally reconstructed \*ī and \*ū can be reinterpreted as \*iH and \*uH in almost all cases, while the only real instances of long \*ī and \*ū appear in cases of monosyllabic lengthening (p. 54–55). Generally speaking, long \*ī and \*ū remain unchanged in IE languages. Long *ī* is written as <y> in Lithuanian, as <ei> in Gothic, and as <i> in Old Irish. Old \*ū changes to *y* in OCS (most likely a *ui*-type diphthong originally – thus written as a combination of *ъ* and *i* in Glagolitic/Cyrillic script), and to *i* in Albanian in final/only syllable (probably *y* otherwise). In Germanic, Italic, and Celtic, \*ī and \*ū shorten in pretonic position in many cases (**Dybo’s Law**). Distinctive length is lost in Armenian, Albanian, and Tocharian. According to some linguists (cf. p. 431), in some cases in Greek (and Tocharian and Armenian) there is a so-called **laryngeal breaking** of \*i/uh<sub>2/3</sub>, where it is not \*i/u that is syllabic but the laryngeals (\*y/w<sub>h</sub>2/3) – cf. Gr. ζω(φ)ός ‘alive’ (⇒ Eng. *zoo-logy*) from \*g<sup>w</sup>yh<sub>3</sub>wos (if not from \*g<sup>w</sup>yeh<sub>3</sub>-?). The reflexes in Greek are thus \*y (palatalization)/φ + ā > η or ω (also Toch. *yā, wā*).

TABLE 1.6 THE REFLEXES OF PIE \*CRHC

PIE	OInd.	Av.	Gr.	Lat.	OCS	Lith.	Goth.	OIr.	Arm.	Alb.	Toch.
*ṛH	<i>ā</i>	<i>ā</i>	μη/μᾱ > μη/μω, ἐμε/ᾰμα/ᾰμο	<i>mā</i>	<i>ę, ɔ</i>	<i>im/ūm</i>	<i>um</i>	<i>má</i>	<i>ama</i>		<i>mā</i>
*ṛ̥H	<i>ā</i>	<i>ā</i>	νη/νᾱ > νη/νω, ἐνε/ᾰνα/ᾰνο	<i>nā</i>	<i>ę, ɔ</i>	<i>in/ūn</i>	<i>un</i>	<i>ná</i>	<i>ana</i>		<i>nā</i>
*l̥H	<i>īr, ūr</i>	<i>ar<sup>o</sup></i>	λη/λᾱ > λη/λω, ἐλε/ᾰλα/ᾰλο	<i>lā</i>	<i>l̥, l̥b</i>	<i>il/ūl</i>	<i>ul</i>	<i>lá</i>	<i>ala</i>	<i>al, la</i>	<i>lā</i>
*l̥̄H	<i>īr, ūr</i>	<i>ar<sup>o</sup></i>	ρη/ρᾱ /ρω, ἐρε/ᾰρα/ᾰρο	<i>rā</i>	<i>r̥, r̥b</i>	<i>ir/ūr</i>	<i>aúr</i>	<i>rá</i>	<i>ara</i>	<i>ar, ra</i>	<i>rā</i>
*iH	<i>ī</i>	<i>ī</i>	<i>ī</i>	<i>ī</i>	<i>i</i>	<i>ý</i>	<i>ī</i>	<i>í</i>	<i>i</i>	<i>i</i>	<i>'i</i>
*uH	<i>ū</i>	<i>ū</i>	<i>ū</i>	<i>ū</i>	<i>y</i>	<i>ú</i>	<i>ū</i>	<i>ú</i>	<i>u</i>	<i>i, y</i>	<i>u</i>

\*CRHV (pre-vocalic position)

In this position, syllabic resonants are not in the same syllable with laryngeals, so this is not really the case of “long syllabic resonants”, and reflexes in most languages are distinct from both \*CRHC and \*CR̥C types:

TABLE 1.7 THE REFLEXES OF PIE \*CRHV

PIE	OInd.	Av.	Gr.	Lat.	OCS	Lith.	Goth.	OIr.	Arm.	Toch.
*ṛH	<i>am</i>	<i>am</i>	αμ <i>em</i>	<i>em</i>	<i>ɛm/ɛm</i>	<i>im/um</i>	<i>um</i>	<i>am</i>	<i>am</i>	<i>ām/am</i>
*ṛ̥H	<i>an</i>	<i>an</i>	αν <i>en</i>	<i>en</i>	<i>ɛn/ɛn</i>	<i>in/un</i>	<i>un</i>	<i>an</i>	<i>an</i>	<i>ān/an</i>
*l̥H	<i>ir/ur</i>	<i>ar</i>	αλ <i>al</i>	<i>al</i>	<i>ɛl/ɛl</i>	<i>il/ul</i>	<i>ul</i>	<i>al</i>	<i>al</i>	<i>āl/al</i>
*l̥̄H	<i>ir/ur</i>	<i>ar</i>	αρ <i>ar</i>	<i>ar</i>	<i>ɛr/ɛr</i>	<i>ir/ur</i>	<i>aúr</i>	<i>ar</i>	<i>ar</i>	<i>ār/ar</i>

Cf. PIE \*ṛ̥h<sub>2</sub>us (p. 16), \*g<sup>w</sup>ṛ̥H- (p. 25) or PIE \*g<sup>w</sup>ṛ̥h<sub>2</sub>us ‘heavy’ > Ved. *gurús*, Gr. βαρύς, Goth. *kaúrus*. The *-ur-* in Ved. *gurús* (cf. also *ūr* after labials – p. 38) is often taken as a trace of the old labiovelar in Ilr., though there are counterexamples (cf. *giris* – p. 25).

PIE VOWELS

PIE had the following vowels (not counting the diphthongs): \*e, \*ē, \*o, \*ō, \*a. The phonemes \*i and \*u were also vowels but phonologically speaking, at least originally, syllabic glides (p. 30–31) (phonetically speaking, *j/i* and *w/u* are generally the same except for the difference in tongue hight/syllabicity). Vowel \*a was a marginal phoneme in PIE (p. 42), with a position somewhat similar to that of \*b (p. 16–17). Some IE-ists even believe that there is no need to reconstruct a PIE \*a. What seems to point to long \*ā is actually always \*eh<sub>2</sub> (p. 45). PIE long \*ē and \*ō were rather rare, limited to certain morphological (p. 61–62) and derivational forms and contractions (p. 65). The reflexes adduced below are sometimes relevant only for non-final and/or accented syllables (e.g., in Latin).

\*e

PIE \*(h<sub>2</sub>)nepōts ‘grandson, nephew’ > Ved. *nāpāt* ‘grandson’ (YAv. *napā* ‘grandson’), Gr. ἀνεπιός ‘cousin’, νεποδες ‘descendants’, Lat. *nepōs* ‘grandson, descendant’ (⇒ Eng. *nepotism*), CS *netii* ‘nephew’, OLith. *nepuotis* ‘grandson’, OEng. *nefa* ‘grandson, nephew’ (Eng. *nephew* ← French), OIr. *necht* ‘niece’, Alb. *nip* ‘grandson, nephew’, *mbesë* ‘granddaughter, niece’

**PIE \*der-** ‘tear, flay’ > Ved. *dar-* ‘split, blow up’, Gr. δέρω ‘I flay’ (δέρμα ‘(flayed) skin’ ⇒ Eng. *dermatology*), OCS *dero* ‘I flay’ (Slav. also ‘tear’), Lith. dial. *derù* ‘I flay’, Goth. *dis-tairan* ‘to tear asunder’ (Eng. *tear*), Arm. *terem* ‘I flay’, Toch. AB *tsār-* ‘be separated’

**PIE \*b<sup>h</sup>er-** ‘bear, carry’ > Ved. *bhárati* ‘bears, brings, keeps’ (YAv. *bara’ti* ‘brings’), Gr. φέρω ‘I bear, carry’, Lat. *ferō* ‘I bear, carry’, OCS *berō* ‘I gather, select’, Goth. *baíran* ‘to bear, carry’ (Eng. *bear*), OIr. *berid* ‘carries, brings’, Arm. *berem* ‘I bear, bring’, Alb. *bie* ‘carry, bear, bring’, Phryg. αἴββερετ ‘brings’, Toch. AB *pār-* ‘bear (away), carry (off)’

PIE \*e remains unchanged in Greek, Latin (in the formerly stressed first syllable, where there is no vowel reduction – p. 328), Balto-Slavic, Celtic, Armenian, and Anatolian (cf. Hitt. *nēpiš* with lengthening under accent – p. 31), except in specific conditions. In Indo-Iranian, PIE \*e causes palatalization (p. 206) and then merges with \*o and \*a into IIr. *a*. In Germanic, PIE \*e yields \*e or \*i (p. 395) – in Gothic it is always *i* (cf. Goth. *itan* ‘to eat’ but ONor. *eta* ‘to eat’, Lat. *edō* ‘I eat’ < PIE \*h<sub>1</sub>ed-), merging with PIE \*i (the new *i* changes to *ai* [ε] before *h, hv, r*). In Albanian the reflexes (*e, ja, ie, je, i*) depend on the phonetic surroundings but are not completely clear (p. 561). PIE \*e merges with \*i into *ä* in Tocharian (p. 455).

## \*o

**PIE \*potis** ‘lord, husband’ > Ved. *pátis, páty-* (GAv. *pá’ti-*), Gr. πόσις ‘husband’, Lat. *potis* ‘(cap)able’, OCS *gos-podъ* ‘lord’, Lith. *pàts* ‘husband, self’, Goth. *bruþ-faþs* ‘bridegroom’, Toch. B *pets* ‘husband’

**PIE \*domh<sub>2</sub>os** ‘house, home’ > Ved. *dámas* ‘house’, Gr. δόμος ‘house’, Lat. *domus*, OCS *domъ*, Lith. *nāmas* ‘house’ < \*damas (assimilation)

**PIE \*g<sup>h</sup>ostis** ‘stranger, guest’ > Lat. *hostis* ‘stranger, enemy’ (⇒ Eng. *hostile*), OCS dat. pl. *gostemъ* ‘to guests’, Goth. *gasts* ‘stranger, guest’ (Eng. *guest* ← ONor. *gestr*), (?) Lep. *Uvamo-kozis* (‘having the highest guests’ (pers. name)), (?) Luw. *kaši-* ‘visit’

IE languages can be divided into two groups when it comes to the reflex of PIE \*o. In one group, PIE \*o generally remains *o* – Greek, Latin (in the first syllable – p. 328), Celtic, Armenian (in most cases, p. 424), Lusitanian, etc. In the other languages, PIE \*o yields *a* (and thus merges with the reflex of PIE \*a/h<sub>2</sub>e) – in Indo-Iranian (\*e/o > *a*), Balto-Slavic (in Slavic, BSl. \*a shifts back to *o*), Germanic, most of Anatolian (including Hittite, but not in Proto-Anatolian and all Anatolian languages – p. 176), Albanian (cf. *natë*, p. 32), Messapian, etc. In Indo-Iranian, PIE \*o > *ā* in open syllables (**Brugmann’s Law**), cf. Ved. *dāru* (p. 17). In Tocharian, PIE \*o yields *e* (Toch. B), *a* (Toch. A), and sometimes *ä* (p. 455).

## \*ē

**PIE \*meh<sub>2</sub>tēr** ‘mother’ > Ved. *mātā* (GAv. *ptā* < \*ph<sub>2</sub>tēr, p. 46), Gr. μήτηρ, Lat. *māter* < \*mātēr (⇒ Eng. *maternal*), Lith. arch. *mótė*, OIr. *máthir* (< PCelt. \*māīr), Toch. B *mācer*

**PIE \*wēg<sup>h</sup>s-** ‘conveyed’ (cf. p. 33, 96) > Ved. *á-vāksur* ‘they drove’, Lat. *uēxī* ‘I carried, conveyed’, OCS *věsъ* ‘I conveyed’

**PIE \*dyēm** ‘day sky’ (acc., cf. p. 56, 70, 72) > Ved. *dyām*, Gr. Ζήν ‘Zeus’, Lat. *diem* ‘day’ (with shortening)

In general, PIE \*ē reflects just like PIE \*eh<sub>1</sub> (p. 45). PIE \*ē is preserved in Greek (<η>), Latin (shortened outside of the first syllable), Hittite (shortened when not accented), and Germanic. In Lithuanian (closed é, p. 488) and Slavic (\*ě, p. 522), it yields a type of long ē. In Indo-Iranian, PIE \*ē (like \*ō and \*ā) yields ā. PIE \*ē changes into \*ī in Celtic and *i* in Armenian. In Albanian the reflex is *o* (p. 561), and in Tocharian PIE \*ē merges with \*o into *e* (Toch. B)/*a* (Toch. A), but only old \*ē causes palatalization (p. 455).

**\*ō**

- PIE \*k<sub>w</sub>ō(n)** ‘dog’ > Ved. *śvā* (YAv. *spā*), Gr. κῶων, Lith. *šuō*, OIr. *cú*, Arm. *šun*
- PIE \*h<sub>2</sub>ekmō(n)** ‘stone’ > Ved. *ásmā*, Gr. ἄκμων ‘anvil, meteoric stone’, Lith. *akmuō*
- PIE \*h<sub>2</sub>ōwyom** ‘egg’ > Pers. *xāya*, Gr. ᾠόν, Lat. *ōuum* (⇒ Eng. *ovulation*), PSI. \*āje (> BCMS *jāje*), Arm. *ju* (with an unexpected *j*-), OW *ui* (< PCelt. \*āwyo-), Alb. *ve*

In general, PIE \*ō reflects just like PIE \*eh<sub>3</sub> or \*oH (p. 46). PIE \*ō is preserved in Greek (<ω>), Latin, and Germanic (cf. Goth. *fotus*, p. 15). Proto-Balto-Slavic \*ō yields *uo* in Lithuanian/Latvian and *a* in Slavic. In Indo-Iranian, PIE \*ō (like \*ē and \*ā) yields ā. In Celtic, PIE \*ō gives \*ā in non-final (OIr. <á>) and \*ū in final/only syllables. Proto-Anatolian \*ō yields Hitt. *ā* (accented) or *a* (unaccented), cf. Hitt. pl. *widār* ‘water’ < \*wedōr (p. 179). In Armenian the reflex of PIE \*ō is *u*, and in Albanian it is *e* (p. 561). In Tocharian PIE \*ō merges with \*ā and \*Ḥ into ā (p. 455).

**\*a**

- PIE \*dap(n)-** ‘sacrificial meal, feast’ > Gr. δάπτω ‘I devour, consume’, δαπάνη ‘cost, expenditure’, Lat. *daps* ‘sacrificial meal, feast’, ONor. *tafn* ‘sacrifice’, Hitt. *tap-pala-* ‘person responsible for court cooking’, Arm. *tawn* ‘feast’, Toch. B *tāpp-* ‘eat’
- PIE \*kapros** ‘he-goat, boar’ > Ved. *kāpr̥th-* ‘penis’, Gr. κάπρος ‘(wild) boar’, Lat. *caper* ‘he-goat, buck’, ONor. *hafir* ‘he-goat’, W *caer-iwrch* ‘roebuck’
- PIE \*bak-** ‘staff, stick’ > Gr. βάκτρον (⇒ Eng. *bacterium*), Lat. *baculum*, (?) Lith. *baksnóti* ‘to poke’, OIr. *bacc* ‘billhook, angle, bend’

A separate reflex of PIE \*a (preserved almost everywhere as *a*) can be seen only in IE languages where there is no \*o > *a* change, i.e., no merger of PIE \*a and \*o. Languages/branches without merger are Greek, Italic, Celtic, Armenian, Tocharian, and originally Anatolian (though merger occurs later in some languages). In Tocharian, PIE \*a merges with PIE \*ō and \*Ḥ into Toch. ā. In Indo-Iranian, Balto-Slavic, Germanic, Albanian, and some Anatolian languages (like Hittite), PIE \*a and \*o merge into *a*.

**TABLE 1.8 THE REFLEXES OF PIE VOWELS**

PIE	OInd.	Av.	Gr.	Lat.	OCS	Lith.	Goth.	OIr.	Arm.	Alb.	Toch.
*e	<i>a</i>	<i>a</i>	ε	<i>e</i>	<i>e</i>	<i>e</i>	<i>i</i>	<i>e</i>	<i>e</i>	<i>eljaljeljeli</i>	’ā
*o	<i>a</i>	<i>a</i>	ο	<i>o</i>	<i>o</i>	<i>a</i>	<i>a</i>	<i>o</i>	<i>o</i>	<i>a</i>	A <i>a</i> , B <i>e</i>
*ē	ā	ā	η	ē	ě	ė	<i>e</i>	<i>i</i>	<i>i</i>	<i>o</i>	A ’ <i>a</i> , B ’ <i>e</i>
*ō	ā	ā	ω	ō	<i>a</i>	<i>uo</i>	<i>o</i>	á	<i>u</i>	<i>e</i>	ā
*a	<i>a</i>	<i>a</i>	α	<i>a</i>	<i>o</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>	ā
*ā	ā	ā	ᾱ > η	ā	<i>a</i>	<i>o</i>	<i>o</i>	á	<i>a</i>	<i>o</i>	A <i>a</i> , B <i>o</i>

### The problem of PIE \*a (and \*ay, \*aw)

In some cases where the classical IE linguistics reconstructed PIE \*a, one now reconstructs \*h<sub>2</sub>e (p. 44), e.g., 1 sg. perf. \*-h<sub>2</sub>e (p. 98) (cf. also the voc. sg. \*-e(h<sub>2</sub>), p. 66–67). The sequence \*h<sub>2</sub>e- is now almost always reconstructed when it comes to initial \*a-, cf. \*h<sub>2</sub>ekmō(n) (p. 41), even if the laryngeal is not strictly provable – many believe that PIE roots, or at least nouns and verbs, always had consonantal onset, and thus a laryngeal is reconstructed on structural grounds. In the laryngealist reconstruction of PIE, \*a is often reduced to a rather infrequent and slightly suspicious peripheral vowel. Some IE-ists even deny its existence (except as an allophone and perhaps in some loanwords) in PIE in general (cf. Lubotsky 1989), though this appears to be too extreme. The position of PIE \*a (and the diphthongs \*ay and \*aw, p. 50–51) is somewhat similar to PIE \*b (which is typologically strange in both cases) – it is a marginal phoneme (though this depends on how one goes on to reconstruct a number of specific roots), appearing often in loanwords and in dialectal words (attested only in a couple of branches), and often in specific phonological contexts (e.g., near velars, cf. \*kap- ‘take’ > Lat. *capīō* ‘I take’, Goth. *haban* ‘to have’ or \*kan- ‘sing’ > Lat. *canō* ‘I sing’, OIr. *canaid* ‘sings’). PIE \*a, \*ay, and \*aw never appear in grammatical morphemes/affixes and are not a part of the usual PIE ablaut system (p. 61). However, there are a couple of forms where an \*a/ā quantitative alternation is traditionally reconstructed: cf. PIE \*nas- ‘nose’ (Ved. gen. du. *nasós*, CS *nosъ*, OEng. *nasu*) and \*nās- (Ved. nom. du. *nāsā*, Lat. *nāsus*, Lith. *nósis*), and PIE \*sal- ‘salt’ (Ved. *salilám* ‘sea’, Gr. ἄλς, Lat. gen. sg. *salis*, OCS *solb*) and \*sāl- (Lat. *sāl*, Latv. *sāl*). Some would now prefer to reconstruct a \*neh<sub>2</sub>s-/\*nh<sub>2</sub>es- (though one could expect \*\*nh<sub>2</sub>es- here) and \*seh<sub>2</sub>l-/\*sh<sub>2</sub>el- type alternation here (\*nh<sub>2</sub>s- and \*sh<sub>2</sub>l- do not work for Old Indic, which would yield \*\*nis-/\*\*sil-, p. 47) or the like (OInd. *nas-* could also be explained as secondary). Possible instances of \*e/a/ā are few, cf. \*h<sub>2</sub>wes- (Hitt. *huišzi* ‘lives’, Ved. *vásati* ‘dwells’, Goth. *wisan* ‘to be’), which can hardly be unrelated to \*h<sub>2</sub>wastu- (Gr. ἄστυ ‘town’) and (?) \*h<sub>2</sub>wāstu- (Ved. *vāstu* ‘residence, house’) – a reconstruction \*h<sub>2</sub>weh<sub>2</sub>stu-/h<sub>2</sub>wh<sub>2</sub>stu- would be difficult to relate to verbal \*h<sub>2</sub>wes- (though \*e/a/ā “ablaut” is, of course, also problematic).<sup>10</sup> Theoretically, when the word is not attested in Old Indic (where syllabic \*ḥ yields *i*), instead of \*a one could reconstruct \*dh<sub>2</sub>p-, \*bh<sub>2</sub>k-, \*kh<sub>2</sub>p-, etc., but this is hardly convincing in all cases – why would it always be \*h<sub>2</sub> (needed to account for Greek α, p. 47) and why would there be no ablaut in these words? In spite of differing opinions, one can conclude that \*a (and perhaps \*ā) did exist as an infrequent phoneme in the final phase of PIE. However, it is also quite possible that pre-PIE had no phoneme \*a (nor \*ay, \*aw).

### The problem of the PIE vowel system

Traditionally, the PIE vocalic system was reconstructed as having the usual \*ā, \*ē, \*ī, \*ō, \*ū (plus diphthongs). In any case, at least phonetically speaking, the vowel system of PIE in its last phase was not problematic. However, taking laryngeals into account and disregarding marginal phonemes (\*a), syllabic glides (\*i, \*u), diphthongs, and length leaves us with \*e and \*o as the only phonological vowels. This kind of analysis would be typologically problematic, since rare two-vowel languages (like Abkhaz) always have /ə/ and /a/ as phonemes, and three-vowel ones have /a/, /i/, and /u/ (like Classical Arabic), though they always have more allophones. To further complicate matters, the role of \*e and \*o in PIE is also suspicious, since \*e is much more frequent and they are part of the same ablaut pattern (p. 61); i.e., they are mostly dependent on morphological and derivational

conditions and are rarely, if ever, used to differentiate separate noun/verb roots. Though there are roots occurring with \*o only (like \*potis or \*g<sup>h</sup>ostis – p. 40), \*e is usually thought to be the “basic” vowel in a root (especially verbal ones), which means that all (or almost all) forms that occur with \*o also have an \*a-variant of the morpheme. Although long-range comparison is highly controversial, it is curious that PIE vowels rarely seem to have any kind of role in it (the supposed correspondences are usually found among consonants only). These and similar facts lead some to speculative claims that pre-PIE perhaps had only one phonological vowel (or even none – cf. Lehmann 1993: 138), which is again typologically and theoretically highly problematic.<sup>11</sup>

### Laryngeals and vocalism

We have already seen the consonantal reflexes of PIE laryngeals (p. 30). In most cases, though, laryngeals left traces in vocalism by changing neighboring vowels qualitatively or quantitatively. In some positions in certain languages, laryngeals themselves changed to vowels. Thus, we analyze them in the context of vocalism.

#### *The influence of laryngeals on neighboring vowels*

Laryngeals influence neighboring vowels in two ways:

- a) \*h<sub>2</sub> and \*h<sub>3</sub> change the color of the preceding/following \*e
- b) disappearance of laryngeals in closed syllables lengthens all preceding syllabic segments

The first change is already present in PIE (where the change is still just allophonic), while the second one is a post-PIE process that nonetheless occurs everywhere except for some cases in Anatolian (p. 177). For PIE \*e, the following formulas generally stand:

*h <sub>1</sub> e > (*h <sub>1</sub> e >) *e	*eh <sub>1</sub> C/# > *ē
*h <sub>2</sub> e > *h <sub>2</sub> a > *a	*eh <sub>2</sub> C/# > *ā
*h <sub>3</sub> e > *h <sub>3</sub> o > *o	*eh <sub>3</sub> C/# > *ō

Thus, \*h<sub>2</sub> changes the preceding/following \*e to \*a, and \*h<sub>3</sub> changes it to \*o, while \*h<sub>1</sub> does not influence vowel color. Only \*e (and \*ey/ew) are influenced by \*h<sub>2</sub> and \*h<sub>3</sub>, while other vowels (\*o, \*ē, \*ō + \*i, \*u, and other diphthongs) are not (possible original \*aH/Ha cannot be distinguished from \*eh<sub>2</sub>/h<sub>2</sub>e). Laryngeals that disappear in closed syllables – word-internally (\*-VHC-) and word-finally (\*-VH#, originally perhaps only before a pause or in cases when the following word began with a non-syllabic segment) – lengthen compensatorily all preceding syllabic segments (except already long \*ē, \*ō): \*e, \*o (p. 45–46), \*i, \*u (p. 38), \*ṛ, \*ṝ, \*ṝ̄ (p. 37–38). Laryngeals in intervocalic position (\*VHV) color the surrounding vowels but after that drop without compensatory lengthening (cf. gen. sg. \*-eh<sub>2</sub>es > \*-ah<sub>2</sub>as > \*-aas > \*-ās, p. 66).

#### *\*HV (laryngeal + vowel)*

In certain cases, an initial laryngeal (\*#HV-) can be directly reconstructed (e.g., if it yields #h- in Anatolian, p. 177). In others, one can infer it indirectly (see below). Yet in



some cases, an initial laryngeal is reconstructed because of theoretical/structural reasons only – e.g., in words like \*h<sub>1</sub>ekwos (p. 22). Here, one could as easily reconstruct \*ekwos (because \*e- and \*h<sub>1</sub>e- would reflect the same), but nowadays \*h<sub>1</sub>ekwos is often reconstructed because it is believed that most PIE words began with a consonant (p. 42, 52). The combination \*HV can occur word-internally (\*-C/VHV-) as well.

### \*h<sub>1</sub>e

**PIE \*h<sub>1</sub>esti** ‘is’ > Ved. *ásti* (Av. *asti*), Gr. ἐστί, Lat. *est*, OCS *jestь*, OLith. *ėsti*, Goth. *ist* (Eng. *is*), Hitt. *ēšzi*, OIr. *is*, Arm. *ē*

**PIE \*h<sub>1</sub>edmi** ‘I eat’ > Ved. *ádmi*, Gr. ἔδω, Lat. *edō* (⇒ Eng. *edible*), OCS *jamь* (< \*ědmь, with Winter’s Law lengthening – p. 481), Lith. *ėsti* ‘to eat (for animals)’, Goth. *itan* ‘to eat’ (OEng. *etan* > Eng. *eat*), Hitt. *ēmi*

**PIE \*h<sub>1</sub>el(h<sub>1</sub>)(n)-** ‘deer’ > Gr. ἔλαφος, OCS *jelenь*, Lith. *ėlni(a)s*, OIr. *elit* ‘doe, hind’, Arm. *eln* ‘deer-cow, hind’, Toch. A *yäl* ‘gazelle’

The development of \*h<sub>1</sub>e is basically the same as that of \*e (p. 40). In many cases (like \*h<sub>1</sub>el(h<sub>1</sub>)(n)-), the initial \*h<sub>1</sub>e- is reconstructed on structural grounds alone. However, this is not always the case: in \*h<sub>1</sub>es- and \*h<sub>1</sub>ed- the existence of initial \*h<sub>1</sub>- can be proven. It is proven by the initial *a-* of Hitt. 3 pl. pres. *ašanzi* < \*h<sub>1</sub>sonti (Ved. *sánti*, p. 47, 93), *adanzi* < \*h<sub>1</sub>donti, etc. The initial laryngeal is also indicated by the lengthened negative *a-* in Ved. ptcp. *á-sat-* ‘which is not’ < PIE \*h<sub>1</sub>snt-.

### \*h<sub>2</sub>e

**PIE \*h<sub>2</sub>erh<sub>3</sub>-** ‘crush > plough’ > Gr. ἀρόω ‘I plough’, Lat. *arō* ‘I plough’, OCS *orati* ‘to plough’, Lith. *arti* ‘to plough’, Goth. *arjan* ‘to plough’, OIr. *airid* ‘ploughs’, Hitt. *harrā-* ‘grind, crush’, Arm. (*h*)*arawr* ‘plough’, Toch. B *āre* ‘plough’

**PIE \*h<sub>2</sub>eks-** ‘axle, axis’ > Ved. *ákṣas* (YAv. *aša-* ‘shoulder’), Gr. ἄξων, Lat. *axis*, OCS *osь*, Lith. *ašis*, OHG *ahsa*, MW *echel* (OIr. *ais* ‘back’)

**PIE \*h<sub>2</sub>ek(r)-** ‘sharp’ > Ved. *ásris* ‘sharp side, corner’, Gr. ἄκρος ‘topmost, outermost’, Lat. *acūtus* (⇒ Eng. *acute*), OCS *ostrь*, Lith. *astrūs*, OEng. *awel* (Eng. *awl*), Gaul. *Axro-talus* (‘with high forehead’ (pers. name)), Arm. *aseln* ‘needle’, Toch. B *akwatse*

Although \*h<sub>2</sub>e is usually written for (mor)phonological reasons, this was most likely phonetically \*h<sub>2</sub>a already in the last phase of PIE. PIE \*h<sub>2</sub>e reflects just like PIE \*a (p. 41), except for Anat. *ha-* (*-ha-*) and perhaps Arm. *ha-* (p. 430). Initial \*a- is often reconstructed as original \*h<sub>2</sub>e- (p. 42), even if there are no concrete traces of a laryngeal (as in \*h<sub>2</sub>eks- or \*h<sub>2</sub>ek(r)-).

### \*h<sub>3</sub>e

**PIE \*h<sub>3</sub>ekw-** ‘eye’ > Ved. *ákṣi-* (Av. *aš-*), Gr. ὄσσε ‘two eyes’, Lat. *oculus*, OCS *oko*, Lith. *akis*, Arm. *akn*, Toch. B *ek*

**PIE \*h<sub>3</sub>ewis** ‘sheep’ > Ved. *ávis*, Gr. (Hom.) ὄϊς, Lat. *ouis*, OCS *ovьca*, Lith. *avis*, Goth. *awistr* ‘sheepfold’ (Eng. *ewe*), Mlr. *oi*, Hitt./CLuw. *hāwi-*, Arm. *hoviw* ‘shepherd’, Toch. B *awi* ‘ewes’

**PIE \*h<sub>3</sub>ek<sup>t</sup>ō(w)** ‘eight’ > Ved. *aṣṭáu* (YAv. *ašta*), Gr. ὀκτώ (⇒ Eng. *octo-pus*), Lat. *octō* (⇒ Eng. *October*), OCS *osmb*, Lith. *aštuoni*, Goth. *ahtau* (Eng. *eight*), OIr. *ochr<sup>n</sup>* (Gaul. *oxtumetos* ‘eighth’), Toch. B *okt*

PIE \*h<sub>3</sub>e reflects just like PIE \*o (p. 40), except for Anat. initial *ha-* and Arm. *ho-* (p. 430). Alternatively, instead of \*h<sub>3</sub>e, it is possible to reconstruct \*Ho (\*h<sub>2</sub>o- when we have Anat. *ha-*, p. 177). It is impossible to distinguish original \*Ho from \*h<sub>3</sub>e.

*\*VH (vowel + laryngeal)*

In most cases, the reflexes of \*eh<sub>1</sub> and \*eh<sub>3</sub> are identical to the reflexes of \*ē and \*ō (p. 40–41) respectively. Most instances of \*ā seem to originate in \*eh<sub>2</sub> (but cf. p. 42).

### \*eh<sub>1</sub>

**PIE \*meh<sub>1</sub>n(s)-** ‘month; moon’ > Ved. *mās* (GAv. *mā* ‘moon’), Gr. μῆν ‘month’, Lat. *mēnsis* ‘month’, OCS *měšecb*, Lith. *mėnuo* ‘month’, Goth. *mena* (Eng. *moon*), *menoþs* (Eng. *month*), OIr. *mí* ‘month’, Arm. *amis* ‘month’, Toch. B *meñe*

**PIE \*d<sup>h</sup>eh<sub>1</sub>-** ‘put, make, etc.’ > Ved. *dá-dhāmi* ‘I set, make’, Gr. τί-θημι ‘I put, set up’ (⇒ Eng. *thesis*), Lat. *fēcī* ‘I made’, OCS *děti* ‘to do’, Lith. *dėti* ‘to lay/put’, Goth. *ga-deþs* (Eng. *deed*), Hitt. *tēmi* ‘I speak, state’

**PIE \*seh<sub>1</sub>-** ‘sow’ > Lat. *sēmen* ‘seed’ (⇒ Eng. *semen*), OCS *sěti*, Lith. *sėti*, Goth. *mana-seþs* ‘mankind’ (OEng. *sæd* > Eng. *seed*), OIr. *sil* ‘seed’

Only Anatolian differentiates original \*ē from \*eh<sub>1</sub> (p. 176).

### \*eh<sub>2</sub>

**PIE \*meh<sub>2</sub>tēr** ‘mother’ > Ved. *mātā* (Av. *mātar-*), Gr. (Dor.) μάτηρ [ā] (Myc. *ma-te*, Att. μήτηρ), Lat. *māter*, OCS *mati*, Lith. arch. *mótė* (Latv. *māte*), OEng. *mōdor* (Eng. *mother*), OIr. *máthir* (Gaul. dat. pl. *matrebo*), Alb. *motër* ‘sister’ (cf. also \*b<sup>h</sup>reh<sub>2</sub>tēr, p. 18)

**PIE \*peh<sub>2</sub>(s)-** ‘guard, graze’ > Lat. *pāscō* ‘I feed, pasture’, OCS *pasti* ‘to pasture, herd, feed’, Goth. *fodjan* ‘to feed, nourish’ (Eng. *food*), Hitt. *paḫš-* ‘guard’, (?) OIr. *ás* ‘growth’, Toch. B *pāsk-* ‘guard’

**PIE \*k<sup>w</sup>eh<sub>2</sub>s-** ‘cough’ > Ved. *kās-*, CS *kašbljati* ‘to cough’, Lith. *kósėti* ‘to cough’, OEng. *hwōsan* ‘to wheeze’, Alb. *kollë*, Toch. B *kosi*

When comparative evidence points to \*ā, a PIE \*eh<sub>2</sub> (phonetically, and written as such in reconstructions by some – cf. e.g. the chapter on Armenian, \*ah<sub>2</sub> already in PIE) is usually reconstructed now. Traditionally adduced examples for a non-laryngeal \*ā (p. 42) are dubious (though not impossible). In some cases, the \*h<sub>2</sub> in \*eh<sub>2</sub> is substantiated by Anatolian (cf. \*peh<sub>2</sub>(s)-); in others (like \*k<sup>w</sup>eh<sub>2</sub>s-) it cannot be strictly proven (BSl. intonations are not really conclusive in distinguishing long grade and \*VH, as some think – p. 107), and in yet others \*h<sub>2</sub> can be indirectly proven: \*me-h<sub>2</sub>tēr obviously has the same suffix as \*p-h<sub>2</sub>tēr (p. 46), \*b<sup>h</sup>re-h<sub>2</sub>tēr (p. 18), \*d<sup>h</sup>ug-h<sub>2</sub>tēr (p. 47), etc., with the laryngeal being directly attested in Indic as the aspiration of the preceding consonant in the last word (p. 30).

The reconstructed PIE \*eh<sub>2</sub> yields \*ā in most languages (the exception being Hitt. -ah<sub>2</sub>- in some positions). In Indo-Iranian this \*ā merges with \*ē/ō into ā. Latin and Celtic preserve ā, and Armenian has the reflex a. In Greek dialects, ā is also preserved, but it changes into η in Attic-Ionic (but not after ε, ι, ρ in Attic, p. 295). BSl. \*ā is preserved in Slav./Latv. a but changes into o [ō] in Lithuanian. The reflex o is seen also in Albanian and Germanic (PGmc \*ō > Goth. <o> [ō], OEng. <ó> [ō]). In Tocharian, PIE \*ā yields o (Toch. B) and a (Toch. A).

### \*eh<sub>3</sub>

**PIE \*deh<sub>3</sub>-** ‘give’ > Ved. *dá-dāmi* ‘I give’ (GAv. *da-dāt* ‘gives’), Gr. *δί-δομι* ‘I give’, Lat. *dōnum* ‘gift’ (⇒ Eng. *donate*), OCS *dati* ‘to give’, Lith. *dúoti* ‘to give’, OIr. *dán* ‘gift’, Hitt. *dā-* ‘take’, Arm. *tur* ‘gift’

**PIE \*ǵneh<sub>3</sub>-** ‘know’ > Ved. *jñā-* (GAv. *-zān-*), Gr. *γνω-* (if not from \*ǵnh<sub>3</sub>-), Lat. (*g*) *nōscō* ‘I get to know’, OCS *znati* ‘to know’, Lith. *žinóti* ‘to know’, Toch. A *knān-*

**PIE \*peh<sub>3</sub>-** ‘drink’ > Ved. *a-pāt* ‘he drank’, Gr. *πῶμα* ‘beverage’, Lat. *pōtus* ‘drunk’, Lith. *puotà* ‘feast, banquet’, Hitt. *pāš-* ‘swallow’

There is no difference between reflexes of \*eh<sub>3</sub> (phonetically/allophonically \*oh<sub>3</sub>, already in PIE), \*oH, and \*ō (p. 41).

### \*CHC (interconsonantal laryngeal)

Laryngeals between consonants (\*CHC), in postconsonantal position word-finally (\*-CH#), and in word-initial pre-consonantal position (\*HC-) are called “syllabic” and thus sometimes marked as \*Ḥ (in \*CRHC it is always the resonants that are syllabic except with the present tense nasal infix – p. 31, 39). However, it is quite possible that they were not really phonetically syllabic but rather pronounced with a prop-vowel \*ə.

### \*Ch<sub>1</sub>C

**PIE \*h<sub>2</sub>enh<sub>1</sub>-** ‘breathe’, **\*h<sub>2</sub>enh<sub>1</sub>mos** ‘breath’ > Ved. *ániti* ‘breathes’ (YAv. *āntyā-* ‘inhalation’), Gr. *ἄνεμος* ‘wind’, Lat. *animus* ‘mind, spirit’ (< \*anamos, ⇒ Eng. *animate*), Goth. *uz-anan* ‘to breathe out’, MW *anadl* ‘breath’ < PCelt. \*ana-tlā, Arm. *holm* ‘wind’ (< \*honm), Toch. B *anāsk-* ‘breathe, inhale’

**PIE \*ǵenh<sub>1</sub>tōr** ‘begetter’ > Ved. *jánitā* ‘begetter, father’, Gr. *γενέτωρ* ‘begetter, ancestor’, Lat. *genitor* ‘begetter, parent, father’ (< \*genatōr); cf. OCS *zěť* (BCMS *zět*), Lith. *žéntas* ‘son-in-law’

**PIE \*wemh<sub>1</sub>-** ‘vomit’ > Ved. *vam(i)-*, Gr. *ἐμέω* ‘I vomit’, Lat. *uomitus* ‘vomiting’ (⇒ Eng. *vomit*), Lith. *vėmti* ‘to vomit’, Goth. *ga-wammis* ‘unclean, stained’ (< PGmc \*wammaz < \*womh<sub>1</sub>n-)

### \*Ch<sub>2</sub>C

**PIE \*ph<sub>2</sub>tēr** ‘father’ > Ved. *pitā* (GAv. *ptā*, dat. sg. *fādrōi*, YAv. *pita*, OPers. *pitā*), Gr. *πατήρ*, Lat. *pater*, (?) CS *strъi* ‘uncle’ (< \*ptr-), Goth. *fadar* (Eng. *father*), OIr. *ath(a)ir* (Gaul. dat. pl. *atrebo*), Arm. *hayr*, Toch. B *pācer*

**PIE \*d<sup>h</sup>ugh<sub>2</sub>tēr** ‘daughter’ > Ved. *duhitā́* (GAv. *dug<sup>o</sup>dar-*), Gr. θυγάτηρ, Osc. *futír*, OCS *dъšti* (< PSI. \*d<sup>h</sup>ukt<sup>i</sup>), Lith. *duktė́*, Goth. *dauhtar* (Eng. *daughter*), Gaul. *dux-tir* (Cib. gen. sg. *tuateros*), Hitt. *duttariyatil/aš* ‘female functionary’ (≠ Luw.?, HLuw. *tuwatra/i-*), Arm. *dustr*, Toch. B *tkācer*

**PIE \*senh<sub>2</sub>-** ‘(seek to) accomplish, reach’ > Ved. *san(i)-* ‘win, reach, get’, Hitt. *šanḫzi* ‘seeks, looks for, attempts’

### \*Ch<sub>3</sub>C

**PIE \*h<sub>2</sub>erh<sub>3</sub>trom** ‘plough’ (cf. \*h<sub>2</sub>erh<sub>3</sub>-, p. 44) > Gr. ἄροτρον, Lat. *arātrum* (with analogical \*ā after *arāre* ‘to plough’), Cz. *rádl*o, Lith. *árklas* ‘wooden plough’ (Latv. *ar̄kls*), OIr. *arathar* (< PCelt. \*aratro-)

**PIE \*dh<sub>3</sub>-tos** ‘given’ (cf. \*deh<sub>3</sub>-, p. 46) > Ved. *a-di-mahi* ‘we shared’, Gr. δοτός, Lat. *datus* (⇒ Eng. *date*)

**PIE \*ph<sub>3</sub>tos** ‘drunk’ (participle, cf. \*peh<sub>3</sub>-, p. 46) > Gr. ποτός ‘drinkable’

In certain languages and in some conditions laryngeals can drop entirely in this position – cf. the word ‘daughter’ with the vocalic outcome in Greek (as always there), Vedic, Tocharian, and likely Celtiberian, but with the zero-reflex in other languages. The only language that distinguishes \*h<sub>1</sub>, \*h<sub>2</sub>, and \*h<sub>3</sub> in this position is Greek, where one gets ε/α/o respectively. In Old Indic, the reflex is usually *-i-*, and in certain unclear conditions it is *-ī-* (cf. Ved. *bráviṭi* ‘speaks’ < \*mlewh<sub>2</sub>-), and occasionally *-Ø-* (cf. Ved. *dadhmás* ‘we place’ < \*d<sup>h</sup>ed<sup>h</sup>h<sub>1</sub>mes). In Iranian, the laryngeals are much less often vocalized to *i* (p. 267). It is not clear why Old Indic sometimes shows the aspiration of the preceding stop (*duhitā́* ‘daughter’) and sometimes does not (*pitā́* ‘father’) (p. 30). This is perhaps related to a different position of the prop-vowel: \*d<sup>h</sup>ugh<sub>2</sub>[ə]tēr (p. 47) but \*p[ə]h<sub>2</sub>tēr (p. 46). In Tocharian we find *ā*, and in other languages, including Armenian (p. 430), it is *a* (unless dropped altogether). In Balto-Slavic, \*VRHC yields \*VRC with the acute accent (PSI. \*, BCMS. /, Lith. /, Latv. / ^) from the laryngeal. In Anatolian, *h* is preserved in some conditions.

### \*#HC- (initial pre-consonantal laryngeal)

#### \*h<sub>1</sub>C-

**PIE \*h<sub>1</sub>se/onti** ‘they are’ > Ved. *sánti* (GAv. *həntī*), Gr. (Myc.) *e-e-si /ehensi/*, Lat. *sunt*, OCS *sotъ*, Goth. *sind*, Hitt. *ašanzi*, Arm. *en*

**PIE \*h<sub>1</sub>lewd<sup>h</sup>-** ‘grow (up); people; free’ > Ved. *ródhati* ‘grows’ (GAv. *raodəṇti* ‘they grow’), Gr. ἐλεύθερος ‘free’, Lat. *liber* ‘free’ (⇒ Eng. *liberty*), *liberī* ‘children’, OCS *ljudie* ‘people’, Lith. *liáudis* ‘people, nation’, Goth. *liudan* ‘to grow’ (OHG *liut* ‘people’)

**PIE \*h<sub>1</sub>su** ‘good’ > Ved. *sú-* (Av. *hūš-*), Gr. εὖ- (⇒ Eng. *eu-thanasia*), OCS *sъ-* (in *sъ-dranъ* ‘healthy’, *sъ-mrěbъ* ‘death’), OIr. *su-*

#### \*h<sub>2</sub>C-

**PIE \*h<sub>2</sub>stēr** ‘star’ > Ved. instr. pl. *stṛ̥bhis* (Av. *star-*), Gr. ἀστήρ (⇒ Eng. *asteroid*), Lat. *stēlla* (⇒ Eng. *stellar*), Goth. *stairno* (Eng. *star*), OIr. *ser*, Hitt. *ḫašterza /ḫšterz/*, Arm. *astl*, Toch. B *šciryē*

**PIE \*h<sub>2</sub>weh<sub>1</sub>-** ‘blow’ > Ved. *vāti* ‘blows’ (YAv. *vāti*), Gr. ἄημι ‘I blow’, Lat. *uentus* ‘wind’, OCS *vějati* ‘to blow’, Lith. *vėjas* ‘wind’, Goth. *waiian* ‘to blow’, *winds* (Eng. *wind*), MW *gwynt* ‘wind’ (< PCelt. \*wintos), Hitt. *ḫuwanza* /*ḫwanzl* ‘wind’, Toch. A *want*

**PIE \*h<sub>2</sub>nēr** ‘man’ > Ved. *náras* ‘men’ (GAv. *nā*), Gr. ἀνῆρ (⇒ Eng. *Andrew*, *android*), Osc. *niir*, MW *ner* ‘chief, hero’, (?) Hitt. *in-narawawar* ‘strength’, Arm. *ayr*, Alb. *njer(i)* ‘human being’, NPhryg. *αναρ*

### \*h<sub>3</sub>C-

**PIE \*h<sub>3</sub>neh<sub>3</sub>m̥** ‘name’ > Ved. *nāma* (YAv. *nama*), Gr. ὄνομα (⇒ Eng. *onomastics*), Lat. *nōmen* (⇒ Eng. *nominal*), OCS *imę* (< BSl. \*inmen- < \*h<sub>3</sub>nh<sub>3</sub>men-), Goth. *namo* (Eng. *name*), OIr. *ainm* (< PCelt. \*anman < \*h<sub>3</sub>nh<sub>3</sub>men-), Hitt. *lāman*, Arm. *anun*, Toch. B *ñem*, NPhryg. *ονομαυ*

**PIE \*h<sub>3</sub>b<sup>h</sup>ruHs** ‘(eye)brow’ > Ved. *bhrūs* (YAv. dat. pl. *brvatbyqm*), Gr. ὄφρῶς, OCS instr. pl. *brǫvǫmi*, Lith. dial. *bruviš*, OEng. *brū* (Eng. *brow*), OIr. *for-brú*, Toch. B *pārǰāne* ‘brows’, Ancient Macedonian (Hesychius) ὀβροῦτες

**PIE \*h<sub>3</sub>neyd-** ‘revile’ > Ved. *nindanti* ‘they revile’ (YAv. *nāismī* ‘I revile’), Gr. ὄνειδος ‘reproach’, Lith. *niédėti* ‘to scorn, despise’, Goth. *ga-naitjan* ‘to revile’, Arm. *anicanem* ‘I curse’

Initial pre-consonantal laryngeals disappear in most languages. One exception is Anatolian, where \*h<sub>2</sub>-/h<sub>3</sub>- > *h*- in some positions and \*h<sub>1</sub> > *a*- (which is disputed) (p. 47). The other is the Greek-Macedonian-Phrygian-Armenian dialectal group, where \*h<sub>1</sub>, \*h<sub>2</sub>, h<sub>3</sub> vocalize – the reflexes (traditionally called “prothetic vowels”) are *ê*-, *â*-, *ô*- respectively in Greek (and Phrygian) and *e*- (> *i*-), *a*-, *o*- (> *a*-) respectively in Armenian (p. 429–430).

### Diphthongs

In PIE, all combinations of vowels and resonants (even \*er, \*om, etc.) are often regarded as diphthongs (like in modern Lithuanian). However, only the combinations of vowels and \*y/w were diphthongs in a strict sense:

*ēy	*ōy	*ay
*ēw	*ōw	*aw

Only these have special reflexes in IE languages – e.g., PIE \*ey does not reflect as a mere sequence of \*e and \*y, while \*er does (cf. \*wert-, p. 33), so there is no need to treat the latter separately in comparative phonology. Long diphthongs (\*ēy, \*ōy, \*ēw, \*ōw) are rare, just like the long-grade \*ē, \*ō (p. 40–41). The position of \*ay and \*aw is similar to that of \*a (p. 42). Secondary \*a- and \*o-diphthongs can arise from laryngeal combinations (p. 43): \*h<sub>2</sub>ey/w > \*ay/w, \*eh<sub>2</sub>i/u > \*ay/w and \*h<sub>3</sub>ey/w > \*oy/w, \*eh<sub>3</sub>i/u (or \*oHi/u) > \*oy/w. Cf. PIE \*neh<sub>2</sub>us ‘ship’ (Gr. νῆς) with \*eh<sub>2</sub>u > \*a(h<sub>2</sub>)u > \*aw, but the clear presence of the laryngeal in the acc. sg. \*neh<sub>2</sub>w̥m̥ (Gr. νῆ(φ)α, Ved. *nāvam*, Lat. *nāuem*) with \*eh<sub>2</sub>w > \*āw. Only tautosyllabic (pre-consonantal) Vy/wC and (word-final) Vy/w# sequences are diphthongs – heterosyllabic (pre-vocalic) Vy/wV sequences are not. Cf. the special diphthongal reflex of \*ey in PIE \*weyd- ‘know’ (Ved. *ved-*, p. 49) or *i*-stem voc. sg. \*-ey (Ved. *-e*, p. 72), but regular reflex of \*e and \*y in non-diphthongal PIE \*treyes ‘three’ (Ved. *tráyas*).

There are four basic kinds of development of diphthongs in IE languages: (a) an unchanged diphthong (e.g., PIE \*ew > Gr. ευ), (b) a different diphthong (e.g., PIE \*ow >

Gmc \*aw), (c) a monophthong (e.g., PIE \*ey > OCS *i*), and (d) subsequent monophthongizations and diphthongizations (e.g., PIE \*ey > East Baltic \*ē > Lith. *ie*). Both monophthongizations (of diphthongs) and diphthongizations (of long vowels) are typologically common. PIE diphthongs are best preserved in (the oldest stages of) Ancient Greek.

### \*ey

**PIE \*deywos** ‘god’ > Ved. *devás* (GAv. *daeua-* (pre-Zarathustra god), OPers. *daiva-* ‘false god’), Lat. *dīuus* ‘godly’ (OLat. *deiuos* ‘god’; ⇒ Eng. *divine*), (?) OCS *divъ* ‘astonishment’, Lith. *diēvas* ‘god’, *deivė* ‘goddess’ (Latv. *dievs*, OPruss. *deiwa(s)*), OEng. *Tīw* (name of a god, cf. *Tues-day*), OIr. *dia* (Gaul. *Deuo-gnata* ‘born of god’ (pers. name), Clb. *Tēiuo-reikis* ‘god’s king’ (pers. name))

**PIE \*h<sub>1</sub>ey-** ‘go’ > Ved. *éti* ‘goes’ (YAv. *a-ēti*, OPers. *aitiy*), Gr. εἶμι ‘I go’, Lat. *īs* ‘you go’ (Paelignian pl. *eite* ‘go!’), OCS *iti* ‘to go’, Lith. *eiti* ‘to go’ (Latv. *iēt* ‘to go’, OPruss. *ēit* ‘goes’)

**PIE \*weyd-** ‘see; know’ > Ved. *védas* ‘knowledge, insight’, Gr. εἶδον ‘I saw’ (⇒ Eng. *idol*), Lat. *uisus* ‘sight’, OCS *viděti* ‘to see’, Lith. dial. *veizdėti* ‘to look’, Goth. *fair-weitjan* ‘to observe’, OIr. *fiadu* ‘witness’

PIE \*ey remains unchanged in early Greek, Old Latin, Old Prussian, and sometimes in Lithuanian. In Indo-Iranian, where \*e/o/a > \*a, \*ey/oy/ay merge into \*ay (cf. \*ew/ow/aw > \*aw), which is preserved in OPers. *ai*, yields Av. *aē* in open and *ōi* in closed syllables (*ē* finally), and becomes *e* in Old Indic (always long and still considered a “diphthong” in Old Indic grammar tradition). Earliest Greek εἶ /ei/ monophthongized to a closed long /ē/ by the time of the Classical Greek in Athens in the 5th century BCE (and to /ī/ even later) – p. 296. Old Latin still had *ei* (up to Plautus and Terence), which later yielded *ī*. BSL. \*ey was preserved in Old Prussian and in certain conditions in Lithuanian, while in others \*ey > \*ē > *ie* (p. 488; traditionally, it is thought that \*ey/oy/ay yield *ie* originally only in stressed syllables, which would later be obscured by analogies). In Slavic, \*ey > *i* (originally always long). In Germanic, PIE \*ey > \*ī (written in Gothic as <ei>). In Celtic, \*ey > \*ē (perhaps not in Celtiberian and Lepontic, p. 360, 362; OIr. *é* > *ia*). PIE \*ey monophthongizes to PANat. \*ē (Hitt. *e*, p. 176), to *i* in Tocharian and Albanian (p. 562), and to *ē* in Armenian (p. 425).

### \*oy (\*h<sub>3</sub>ey, \*eh<sub>3</sub>i)

**PIE \*loyk<sup>w</sup>(n)-** ‘remaining’ > Ved. *rékṇas* ‘inherited property’ (GAv. *raēx<sup>n</sup>nah-* ‘inheritance’), Gr. λοιπός ‘remaining’, OCS *отъ-лѣкъ* ‘remainder’, Lith. *āt-laikas* ‘remainder’, OHG *lēhan* (< PGmc \*layx<sup>w</sup>naz, Eng. *loan*)

**PIE \*woy<sup>k</sup>-** ‘settlement’ > Ved. *veśás* ‘inhabitant’ (YAv. *vaēsa-* ‘servant’), Gr. οἶκος ‘house’ (⇒ Eng. *eco-nomy*, etc.), Lat. *uīcus* ‘village’ (OLat. *ueicus*), Lith. *viēš-pats* ‘lord’ (OPruss. acc. sg. *wais-pattin* ‘housewife’)

**PIE \*h<sub>1</sub>oy<sup>n</sup>/k/wos** ‘one’ (p. 89) > Ved. *ékas* (Av. *a-ēua-*, OPers. *aiva-*), Gr. οἶος ‘alone’, Lat. *ūnus* (OLat. acc. sg. *oino*), Lith. *vienas* (Latv. *viēns*, OPruss. *ains*), Goth. *ains* (Eng. *one*), OIr. *óen*

PIE \*oy is preserved in Greek (οι), Old Latin, and Celtic (written as <oi>, <oe>, etc. in Old Irish). See \*ey for Indo-Iranian. OLat. *oi* remains a diphthong (written as <oe>) in

seven words after labials (cf. *foedus* ‘treaty’) and yields *ū* elsewhere in Classical Latin. (initial *\*woy-* > *ū-*, however). In Balto-Slavic, *\*oy* and *\*ay* merge into *\*ay*, preserved as *ai* in Old Prussian and yielding, as with *\*ey* above, *ai* in some positions in Lithuanian and *ie* (via *\*ē*) in others (p. 488). In Slavic, the reflex is *ě* and sometimes *-i* finally. In Germanic, *\*oy* and *\*ay* also merge into *\*ay* (Goth. *ai*). Hittite has *ai* for PIE *\*oy* before coronal consonants, otherwise *e* (p. 176). In Armenian, PIE *\*ey* and *\*oy* yield *ē* (p. 425). In Albanian, PIE *\*oy* (and *\*ay*) yield *e* (p. 562). In Tocharian B, it is *ei* (> *ai*), and *e* in Tocharian A.

### **\*ay (\*h<sub>2</sub>ey, \*eh<sub>2</sub>i)**

**PIE \*kaykos** ‘one-eyed, blind’ > Skr. *kekaras* ‘squinting’, Lat. *caecus* ‘blind’, Goth. *haihs* ‘one-eyed’, OIr. *cáech* ‘one-eyed’

**PIE \*laywos** ‘left’ > Gr. *λαίος*, Lat. *laeuus*, OCS *lěvъ*, (?) Toch. B *laiwo* ‘lassitude’

**PIE \*h<sub>2</sub>eydh-** ‘burn’ > Ved. *édhas* ‘firewood’ (YAv. *aēsma-*), Gr. *αἴθω* ‘I kindle’, Lat. *aestus* ‘heat’, Lith. *iesmė* ‘firewood’, OEng. *ād* ‘funeral pile’ (< PGmc *\*ayðaz*), OIr. *áed* ‘fire’, Alb. (*h*)*ethe* ‘fever’

For the problem of *\*a(y)*, cf. p. 42. Initial *\*ay-* is usually interpreted as original *\*h<sub>2</sub>ey-*, while *\*ay* is often assumed after velars (p. 42), though alternative reconstructions like *\*kh<sub>2</sub>eykos* and *\*keh<sub>2</sub>ikos* are not unimaginable. See *\*oy* above for Indo-Iranian, Balto-Slavic, and Germanic. Old Latin still has *ai*, later spelled as <ae> (and pronounced as a diphthong by the upper classes). In Armenian (*ay*) and Celtic, the diphthong is preserved (Old Irish has a late merger with *\*oy*, p. 366). In Hittite, *ai* is preserved before coronals; otherwise it is *e* (p. 176). Tocharian B preserves *ai* (cf. *e* in Tocharian A).

### **\*ew**

**PIE \*b<sup>h</sup>ewdh-** ‘be awake/aware’ > Ved. *bódhati* ‘is aware, notices’ (Av. *baodāti*), Gr. *πεύθομαι* ‘I find out/inquire’, OCS *bljusti* ‘to watch/guard’, Goth. *ana-biudan* ‘to bid/command’

**PIE \*h<sub>1</sub>lewdh-** ‘free; man’ > Gr. *ἐλεύθερος* ‘free’, OCS *ljudie* ‘people’, Lith. *liáudis* ‘people, nation’, OEng. *lēod* ‘man’ (< PGmc *\*lewdōiz*)

**PIE \*h<sub>1</sub>ews-** ‘sing, scorch’ > Ved. *oṣati* ‘singes’, Gr. *εῦω* ‘I singe’, Lat. *ūrō* ‘I scorch/burn’

PIE *\*ew* remains unchanged in Greek (*ew*). In Indo-Iranian, *\*ew/ow/aw* merge into *\*aw* (cf. *\*ey/oy/ay*), preserved in OPers. *au*, yielding Av. *ao* in open and *āu* in closed syllables (*-(au)ō* finally), and *o* in Old Indic (always long and considered a “diphthong”). In Archaic Latin *eu* merges with *ou* < PIE *\*ow* and later yields *ū*. BSL *\*ew* is perhaps preserved in OPruss. *eu* in some cases in Lithuanian it yields *iau* (p. 489, 498), and in Slavic (palatalizing) *iu*. Gmc *\*ew* yields Goth. *iu* (cf. p. 40). In Celtic, *\*ew* tended to change to *\*ow* (> OIr. mostly *úa*, p. 366). Monophthongization occurs in Anatolian (Hitt. *u*, p. 176), Albanian (*e*), and Tocharian (*u*). In Armenian, *\*ew* and *\*ow* merge into *oy*.

### **\*ow**

**PIE \*b<sup>h</sup>owd<sup>h</sup>eye-** ‘wake (somebody) up’ > Ved. *bodháyati* ‘wakes up’ (YAv. *baodaiie-ti* ‘reveals’), OCS *u-buditi* ‘to wake up’, Lith. *báudyti* ‘to incite’

- PIE \*klownis** ‘haunch, thigh, hip, buttock’ > Ved. *śróni-* ‘hip and loins, buttocks’ (YAv. *sraoni-* ‘hip, buttock’), Lat. *clūnis* ‘buttock, haunch’, Lith. *šlaunīs* ‘thigh, hip’, OIc. *hlaun* ‘buttock, haunch’, MW *clun* ‘thigh’ (< PCelt. \*klownis)
- PIE \*h<sub>1</sub>rowd<sup>h</sup>os** ‘red’ > Ved. *lohás* ‘reddish’, Lat. *rūfus* (a Sabellic loanword), Cz. *rudý*, Lith. *raūdas* ‘reddish’, Goth. *rauþs* (Eng. *red*), OIr. *riad* (Gaul. *Roudius* (pers. name))

For Indo-Iranian and Armenian see \*ew. The earliest Greek still had a diphthongal ou, which became a long closed /ō/ and later /ū/ in Attic (p. 296). OLat. *ou* became Classical Latin *ū*. In Balto-Slavic, \*ow and \*aw yielded \*aw (preserved as Lith. *au*), which in turn gave Slav. *u*. Germanic also had merger of \*ow and \*aw (cf. Goth. *au*). PIE \*ow is preserved in Celtic (> OIr. *ó* > *úa*). Hittite has *u* but *au* before coronals (p. 176) as reflexes for \*ow and \*aw, the Albanian reflex is *a*, while Tocharian B has *eu* (later *au*, cf. Toch. A *u*).

**\*aw (\*h<sub>2</sub>ew, \*eh<sub>2</sub>u)**

- PIE \*h<sub>2</sub>ews-** ‘shine’ > Gr. *αὔριον* ‘tomorrow’, Lat. *aurōra* ‘dawn’, *aurum* ‘gold’, OCS *utro* ‘morning’, Lith. *aušrà* ‘dawn’, *áu(k)sas* ‘gold’, ONor. *austr* (Eng. *east*)
- PIE \*h<sub>2</sub>ews-** ‘ear’ > Lat. *auris*, OCS *uxo*, Lith. *ausis*, Goth. *auso* (Eng. *ear*), OIr. *au/ó* (Gaul. *Su-ausia* (pers. name))
- PIE \*h<sub>2</sub>ewg-** ‘grow’ > Ved. *ójas* ‘strength’ (GAv. *aogō*), Gr. *αὔξω* ‘I grow’, Lat. *augeō* ‘I increase’, Lith. *áugti* ‘to grow’, Goth. *aukandei* ‘increasing’

Concerning the PIE status of \*aw, cf. \*a and \*ay (p. 42). Initially, \*h<sub>2</sub>ew- is usually reconstructed, while \*aw occurs, for instance, in some loanwords (cf. \*tawros, p. 57). For Indo-Iranian see \*ew; for Balto-Slavic, Germanic, and Hittite see \*ow. Greek, Latin, and Celtic preserve the original *aw/au* (OIr. *úa/ó*). PIE \*aw is preserved in Armenian (*aw*) and Tocharian B (*au*, cf. Toch. A *o*). In Albanian, the reflex is *a* (p. 562).

*Long diphthongs*

Like \*ē and \*ō, long \*ēy, \*ōy and \*ēw, \*ōw (original and resulting from various contractions) were rather rare and later merged with short diphthongs in most languages. The special reflexes are clearest in Indo-Iranian, where they yield \*āy (OInd. *ai*, word-finally *-ā*; Av. *āi*) and \*āw (OInd. *au*; Av. *āu*) respectively. Cf. PIE \*g<sup>w</sup>ōws (p. 24, 71) and instr. pl. \*-ōys (p. 65, 266). Greek has special reflexes for \*-ēy(-), \*-ōy(-) (p. 289–290, cf. dat. sg. \*-ōy, p. 65).

**TABLE 1.9 THE REFLEXES OF PIE DIPHTHONGS**

PIE	OInd.	Av.	Gr.	Lat.	OCS	Lith.	Goth.	OIr.	Hitt.	Arm.	Alb.	Toch.
*ey	<i>e</i>	<i>aē/ōi</i>	<i>ει</i>	<i>ī</i>	<i>i</i>	<i>ie/ei</i>	<i>ei</i>	<i>ía/é</i>	<i>e</i>	<i>ē</i>	<i>i</i>	<i>i</i>
*oy	<i>e</i>	<i>aē/ōi</i>	<i>οι</i>	<i>oel/ū</i>		<i>ie/ai</i>	<i>ai</i>	<i>ói (óe)</i>	<i>e/ai</i>	<i>ē</i>	<i>e</i>	<i>A e, B ei &gt; ai</i>
*ay	<i>e</i>	<i>aē/ōi</i>	<i>αι</i>	<i>ae</i>	<i>ě</i>	<i>ie/ai</i>	<i>ai</i>	<i>ái (áe)</i>	<i>e/ai</i>	<i>ay</i>	<i>e</i>	<i>A e, B ai</i>
*ew	<i>o</i>	<i>ao/ēu</i>	<i>ευ</i>	<i>ū</i>	<i>u</i>	<i>iau</i>	<i>iu</i>	<i>úa/ó</i>	<i>u</i>	<i>oy</i>	<i>e</i>	<i>u</i>
*ow	<i>o</i>	<i>ao/ēu</i>	<i>ου</i>	<i>ū</i>	<i>u</i>	<i>au</i>	<i>au</i>	<i>úa/ó</i>	<i>u/au</i>	<i>oy</i>	<i>a</i>	<i>A o, B eu &gt; au</i>
*aw	<i>o</i>	<i>ao/ēu</i>	<i>αυ</i>	<i>au</i>	<i>u</i>	<i>au</i>	<i>au</i>	<i>úa/ó</i>	<i>u/au</i>	<i>aw</i>	<i>a</i>	<i>A o, B au</i>



## PHONOTACTICS

### Syllable structure and consonant clusters

PIE had CVC, CV, and VC type syllables – with consonant clusters also, (CC)CVC(CC), etc. As usual, more types of clusters were permitted in medial (heterosyllabic) than in initial or final (homosyllabic) position – for instance, the cluster \*-mb<sup>h</sup>- was possible medially in \*ǵomb<sup>h</sup>os (p. 18), but not initially (where it would be \*\*ṁb<sup>h</sup>-, p. 31) and finally (there was no final \*-b<sup>h</sup>, p. 53). As in living languages, it is sometimes difficult to ascertain the exact syllabification in medial positions – e.g., is it \*h<sub>1</sub>ek-wos or \*h<sub>1</sub>e-kwos (p. 22), \*ne-b<sup>h</sup>os or \*neb<sup>h</sup>-os (p. 31), or both.

It seems that onsetless initial syllables (#VC-) were rare. It is common practice now to reconstruct initial laryngeals even when not strictly provable, though in some cases they can be proven (p. 44). Thus, the frequency of initial VC- depends on whether one prefers to reconstruct \*ekwos or \*h<sub>1</sub>ekwos (which in cases of \*h<sub>2</sub>e- also depends on one's position on the status of PIE \*a – p. 42). It is possible that #VC- was usual in pronouns, e.g., in forms like \*ey 'that' (p. 87), though here one can also reconstruct \*h<sub>1</sub>ey. In any case, onsetless initial syllables were possible with syllabic resonants, like \*ṅ- 'non-' or \*udn- 'water' (in oblique cases of \*wodḡ). In cases with zero grade only, optional initial laryngeals are theoretically reconstructable, cf. \*(H)ish<sub>1/2</sub>ros (Ved. *iṣírás* 'strong, active', Gr. *ιερός* 'holy; strong').

As for single initial consonants, it seems that only \*r- was not permitted (p. 32), so \*Hr- may be reconstructed instead of \*r- even when not strictly provable (e.g., in \*(H)rotHos – p. 32). Initial \*TR- (like \*klewos 'fame' > Ved. *śrávas* 'renown', Gr. *κλέος*) was not possible in case of homorganic stops + nasals (e.g., \*\*pm-, \*tn-, etc.), though this was normal in medial heterosyllabic position (e.g., \*potni<sub>2</sub> 'mistress' > Ved. *pátñī*, Gr. *πάρνια*). Initial \*sC- was usual (cf. \*steyg<sup>h</sup>- (p. 19), \*sneyg<sup>wh</sup>- (p. 25)), and initial \*ks- was also apparently possible (cf. \*ks-, p. 16). Only some cases of (labial + dental/alveolar) \*RR- (with asyllabic first resonant!) were permitted: \*wr-, \*wl-, \*ml- (cf. \*mlewh<sub>2</sub>- > GAv. *mraomī* 'I say'), \*mr- (cf. comp. \*mreg<sup>h</sup>wyōs 'shorter' > Lat. *breuis* 'short'), and \*mn- (cf. \*mneh<sub>2</sub>- 'think' (LIV)). Before other consonants, resonants were always syllabic: \*R̥C- (p. 31). Initial \*CHV- is suspicious, though some would reconstruct it in some cases of PIE \*a (p. 42, 50). In cases like \*ph<sub>2</sub>tēr, laryngeals were "syllabic" (p. 46). In cases of expected \*CHR̥-, a **laryngeal metathesis** occurs (\*CHR̥- > \*CR̥H-). Cf. \*peh<sub>3</sub>- 'drink' (Ved. aor. *apāt* 'he drank', Lat. *pōtus* 'drunk') but \*ph<sub>3</sub>i- (a zero grade of \*peh<sub>3</sub>-y-) > \*pih<sub>3</sub>- (metathesis) > Ved. *pītás* 'drunk', OCS *piti* 'to drink'. Initial \*HC- (frequently with resonants, less with stops and \*s) was permitted, as proven by "prothetic vowels" in some languages (p. 48), but was perhaps "syllabic" in those positions (cf. p. 46–47).

Initial clusters of two stops were infrequent in PIE, but some cases seem to have existed (in spite of the tendency to eliminate such clusters). The initial cluster \*pk- is apparently seen in Av. *fšūmant-* 'wealthy; having cattle' (cf. PIE \*peku 'cattle' > Ved. *pásu*, Goth. *faihu*). In some cases, a non-phonemic "secondary schwa" appears to break a cluster (which yields *a* in Latin and most languages but apparently *ι* in Greek): cf. \*k<sup>w</sup>twor-[k<sup>w</sup>otwor-]<sup>12</sup> 'four' (in oblique cases) > Lat. *quattuor*, Gr. (Hom.) *πίσυρες* (the cluster is attested without a prop-vowel medially in YAv. *ā-xtīrīm* 'four times' < \*k<sup>w</sup>tury-, and initially in Ved. *turíya-* 'fourth' with probably later cluster simplifications); \*ptneh<sub>2</sub>- [pətneh<sub>2</sub>] > Lat. *pandō*, Gr. *πίντημι* 'I spread out' (cf. also Hitt. *taknāš* below). In other cases, original initial clusters had probably been simplified already in PIE, cf. \*kṁptom (p. 34) <

\*d̥k̥mtom ‘hundred’, obviously derived from \*dek̥m(t) ‘ten’ (p. 34, 90). Some all-stop initial clusters (like Gr. *πτερόν* ‘feather’ from PIE \*pet- ‘fly’) are likely post-PIE, and others (like \*pteh<sub>2</sub>k- ‘to duck’ in LIV) are dubious. A number of original dental + velar clusters seem to have existed – e.g., \*tken- ‘hit, injure’, \*d<sup>h</sup>ǵ<sup>h</sup>uH- ‘fish’, \*d<sup>h</sup>ǵ<sup>h</sup>(e)m- ‘earth’, etc. These clusters were rarely preserved (cf. Toch. A *tkam* ‘earth’) – either the first stop drops (Lith. *žūvis* ‘fish’; OCS *zemlja* ‘earth’); a metathesis occurs (\*tken- > \*kten- [k̥p-] > Ved. *kṣanóti* ‘injures’, Gr. κτείνω ‘I kill’; \*d<sup>h</sup>ǵ<sup>h</sup>em- > \*ǵ<sup>h</sup>d<sup>h</sup>em- [ǵ<sup>h</sup>ǵ<sup>h</sup>-] > Ved. *kṣám-*, Gr. χθών ‘earth’); an epenthetic vowel emerges (Hitt. gen. sg. *taknās* ‘earth’) or, rarely, a prothetic vowel appears (Gr. ἰχθύς ‘fish’). In both initial and medial dental + velar clusters, there is often a metathesis (as in \*d<sup>h</sup>ǵ<sup>h</sup>(e)m- > \*ǵ<sup>h</sup>d<sup>h</sup>(e)m-) already in PIE (probably not in all conditions), and the dental stop (\*t/d/d<sup>h</sup>) changes to an allophonic \*p̥/ǵ/ǵ<sup>h</sup>, as traditionally reconstructed, with special reflexes in some languages (OInd. *kṣ*, Lat. *s* but Gr. κτ/χθ – cf. \*h<sub>2</sub>k̥p̥os, p. 35). In these cases (the so-called **thorn-clusters**), we are thus dealing not with all-stop clusters but with stop + fricative clusters (similar to \*ks- or \*THV- above) if the traditional reconstruction is correct.

As for triconsonantal initial clusters, almost all of them had a fricative (\*s/H) as the first or the second element. Those in \*sTR- were the most usual (e.g., PIE \*streyg-/stri(n)g- > Lat. *stringō* ‘I strip off’, OCS *strigō* ‘I cut’), the \*CsC- ones were rare and likely onomatopoeic (cf. \*pster- > Gr. πτύρνυμαι ‘I sneeze’, Lat. *sternuō* ‘I sneeze’; \*ksweyb<sup>h</sup>- ‘swing’ (LIV); \*ksnew- (p. 16)). Possible clusters with initial laryngeal (\*h<sub>2</sub>stēr (p. 47), \*h<sub>1</sub>syeh<sub>1</sub>t (p. 102), \*h<sub>3</sub>b<sup>h</sup>ruHs (p. 48), etc.) perhaps had a “syllabic” laryngeal (i.e., phonetic \*HəCC- or similar). Triconsonantal clusters without \*s/H seem to be “thorn-clusters” – cf. \*d<sup>h</sup>ǵ<sup>h</sup>yes (> \*ǵ<sup>h</sup>d<sup>h</sup>yes) ‘yesterday’ (> Ved. *hyás*, Gr. χθές) and \*pkten- ‘comb’ (Gr. κτεῖς, cf. Lat. *pecten* < \*pekten-). An initial four-consonant onomatopoeic cluster can also possibly be reconstructed: \*sptyewH- (LIV) > Gr. πτύω ‘I spit out’, Goth. *speiwan* ‘to spit’. A number of additional clusters were permitted word-medially, but we cannot adduce them all here. Geminates were possible only in expressive or nursery words like \*atta ‘daddy’ (Gr. ἄτρα ‘father’, Lat. *atta*, Goth. *atta* ‘father’, Hitt. *attas* ‘father’ – cf. also the two \*a phonemes in the word). The \*-ss- on morpheme boundary was probably simplified to \*-s- already in PIE (cf. p. 78 for *s*-stems and p. 93 for \*h<sub>1</sub>e(s)si).

Obviously, not all segments and clusters were allowed word-finally. A PIE word could end in any syllabic segment, except for \*ē, \*ō (apparent final \*-ē/ō always stem from \*-e/oH), marginal \*a/ay/aw (p. 42), and apparently \*-l. As for single consonants, all resonants and fricatives were permitted word-finally but of stops only \*-t/-d (cf. aor. 3 sg. \*-t (p. 96), \*tod ‘that’ > Goth. *þat-a*, Ved. *tát/d*, OCS *to*), and marginally \*ǵ (only in \*h<sub>1</sub>eǵ ‘I’ – p. 82). It is often claimed that \*-t and \*-d are not distinctive in final position. Indeed, in most languages there are no clear reflexes of such an opposition (and in languages like Greek or Old Church Slavic no final *-t/-d* is even possible). However, cf. the already adduced Goth. *þat-a* < PGmc \*þat (Eng. *that*, Germ. *das*) < PIE \*tod, but Goth. *miliþ* ‘honey’ < PIE \*melit (p. 77). The resonants in \*-CR̥ were always syllabic, and the same goes for \*-CH̥, cf. \*g<sup>w</sup>en̥h<sub>2</sub> (p. 24, 66), \*meǵ<sub>2</sub> [-ǵh<sub>2</sub>ǵ] (p. 30) ‘big’ > Ved. *máhi*, Gr. μέγα. A number of final PIE clusters are morphologically reconstructable (all of them ending in \*-s or \*-t), though it is difficult to know whether some of them were phonetically simplified already in PIE. Considering \*-RT type final clusters, \*-nt was frequent (p. 97), and others appeared occasionally, cf. root aor. \*g<sup>w</sup>emt ‘went’ (p. 96) or root noun \*kērd ‘heart’ (Gr. Hom. κῆρ, Hitt. *ker*), though these are usually structurally reconstructed (i.e., the clusters are not actually attested in daughter languages so it is not impossible that, for instance, \*-rd# was simplified already in PIE). All \*-Rs clusters were possible, though in pre-PIE \*-VRs > \*-VR̥ in nom. sg. (p. 70) (cf. also \*-Vs/nh<sub>2</sub> > \*-VR̥s/n in nom. pl. – p. 70).

The following fricative-first clusters are attested: \*-Hs (cf. \*muHs – p. 38), \*-Ht (cf. root aor. \*deh<sub>3</sub>t ‘gave’ – p. 96), and \*-st (cf. inj. \*h<sub>1</sub>est ‘is’ – p. 93, 97). Clusters of \*-Ts type were usual: \*pōds (p. 15), \*h<sub>3</sub>rēgs ‘king’ (> Lat. *rēx*), \*snoyg<sup>wh</sup>s ‘snow’ (p. 25), etc. All-stop final clusters were exceptional, cf. \*glagt [-kt] ‘milk’ (Gr. γάλα, also γάλαγος, Lat. *lac*). As a parallel to \*(s)TR- initial clusters, there were also some final triconsonantal clusters: \*-CCs (\*nog<sup>wh</sup>ts > \*nok<sup>wh</sup>ts (p. 32) > Lat. *nox*, Goth. *nahts*; ptcp. \*-nts (p. 104)) and \*-Cst (cf. sigm. aor. forms like \*yēwgst ‘yoked’ – p. 96).

Not all combinations of stops were permitted (or usual) in PIE roots (in the onset and coda). There were two **root constraints** (or near-constraints since there are some counter-examples): (a) no aspirated and voiceless stops together (i.e., no roots like \*\*ped<sup>h</sup>- or \*\*g<sup>h</sup>et-) and (b) no two voiced stops together (i.e., no roots like \*\*ged-). Constraint (a) does not include \*sT- initial roots like \*steyg<sup>h</sup>- (p. 19) and constraint (b) is sometimes interpreted glottally (p. 18).

### Phonological rules in PIE

We have already mentioned Pinault’s Law (p. 30), laryngeal metathesis (p. 52), the post-velar allophonic \*b/ð/ð<sup>h</sup> (p. 53), and here we shall briefly list some of the other phonological rules of PIE. Some rules will be discussed in chapter 2 “Proto-Indo-European Morphology” (p. 70) and for Bartholomae’s Law cf. p. 229.

#### Dental clusters

In PIE, when two dentals (\*t/d/d<sup>h</sup>) occurred next to each other (on the morpheme border), an \*-s- was automatically inserted between them: \*wid-tos ‘seen/known’ (cf. \*weyd-, p. 49) > \*wid-s-tos [witstos], \*h<sub>2</sub>ed-te! ‘eat! (pl.)’ (cf. \*h<sub>1</sub>ed-, p. 44) > \*h<sub>1</sub>ed-s-te [h<sub>1</sub>etste]. This \*stst was preserved in Anatolian as <zt> (cf. OHitt. *ēzten!*) but changed in various ways in other languages: to -tt- in Indic (Ved. *vittás* ‘found’, *attá!*); to -st- in Iranian (GAv. *vistō* ‘found’), Greek (ἄ-ιστος ‘unseen’), and Balto-Slavic (OCS *věstb* ‘known’); and to -ss- in Italic (Lat. *uīsus* ‘seen’), Celtic (OIr. *ro-fess* ‘is known’), and Germanic (Goth. *un-wiss* ‘uncertain’).

#### Sievers’ Law and Lindeman’s Law

Evidence in some IE languages (like Vedic and Gothic) points to a possible, but disputed, PIE rule (Sievers’ Law) that \*-y- and \*-w- changed to \*-iy- and \*-uw- after a “heavy” syllable (i.e., after \*-VCC- and \*-VC-), perhaps only in final position. Cf. Goth. *asneis* [-is] ‘day-laborer’ < \*asn(i)jaz but *nipjis* ‘relative’ < \*nipjaz. The effects of the law were later blurred by various analogical developments in separate languages. A variant of the same law occurred in monosyllabics (Lindeman’s Law), where the same distribution (\*-y-/iy-, \*-w-/uw-) would have originally depended on the ending of the preceding word, cf. \*k(u)wō(n) (p. 41) or \*d(u)woh<sub>1</sub> (p. 89). Again, later languages can generalize one or the other form – cf. both in Ved. *ś(u)vā* (only Vedic has clear traces of the law) but only \*-uw- in Gr. κύων (though this can also be by analogy to gen. sg. κυνός).

#### Monosyllabic lengthening

There was a tendency in PIE for monosyllabic forms to lengthen phonetically.<sup>13</sup> Cf. \*ne ‘no(t)’ (Ved. *ná*, Lat. *ne-que* ‘and not’, OCS *ne*, Lith. *nė*, Goth. *ni*) and \*nē (Lat. *nē*, OCS *ně-kako* ‘somehow’, Lith. *nė*, Goth. *ne*, OIr. *ní*); \*nū(n) ‘now’ (p. 81); pers. pronouns \*tū

‘thou’, \*mē ‘me’, \*nōs ‘us’, etc. (p. 82–84); \*wis (or oblique \*wis-) ‘poison’ (Ved. *viśám*) and \*wīs (YAv. *vīš*, Lat. *uīrus* ⇒ Eng. *virus*), etc. This lengthening, exhibited in numerous monosyllabic forms attested in IE languages both with brevity and length, appears in a few nouns with \*i/u (though in \*wīs it may be from older \*wis-s as well – p. 70, 79) and otherwise mostly in personal pronouns and particle-like words (particles, conjunctions, postpositions – cf. e.g., \*prō and \*-wē ‘or’, p. 105), apparently in both accented and unaccented (clitic) forms. Not all monosyllables exhibit both variants, cf. only short \*h<sub>1</sub>eg ‘I’ (p. 82), \*so ‘this’ (p. 85), or \*yo ‘which’ (p. 88), which is in some cases probably due to chance preservance of just one variant and in others probably to various morphological analogies. Monosyllabic lengthening is the only non-laryngeal source of long \*ī and \*ū (p. 38). It is not clear if it originally had anything to do with the stable (and morpho-logical) length in nouns like \*pōds (p. 15).

## ACCENT

Some IE languages have a fixed accent place (like Latin – p. 328), while others have a free accent, i.e., an unpredictable accentual position. Since those that have a free accent (like Vedic and Greek) display certain correspondences, these free accent systems are taken to be inherited from PIE. The following branches/languages preserve (to different extents) the old PIE free accent or traces of it: Vedic (but not Classical Sanskrit – p. 222–223), Avestan (indirectly in marginal traces – p. 272–273), certain modern Iranian (Pashto, perhaps Yidgha) and Dardic languages (Shina, Dameli), Greek (p. 296–297), Balto-Slavic (Lithuanian, Latvian (p. 490–494); Proto-Slavic and certain Slavic languages – East Slavic, Slovene, Bosnian/Croatian/Montenegrin/Serbian, Bulgarian, etc. (p. 523–525)), Germanic (not directly but through Grimm’s and Verner’s Laws – p. 20, 393), and Anatolian (not directly but through vowel lengthening – p. 177).

Traditionally, the reconstruction of the PIE accent is very simple – one compares the accent position of Vedic and Greek, plus indirect information that can be gathered from Germanic consonantism. The data from Anatolian is sparse, and that from Balto-Slavic is rarely used because it is considered to be innovative. When the accent position in Vedic, Greek, and Germanic corresponds (which is often the case), the same stress position is reconstructed for PIE (marked with \*):

Ved. *bhrátā*, Gr. φράτηρ, Goth. *broþar* < PIE \*b<sup>h</sup>rēh<sub>2</sub>tēr ‘brother’ (p. 18)

Ved. *nábhas*, Gr. νέφος (cf. also Hitt. *nēpiš* with lengthening of the accented syllable and OLith. *dēbesis*) < PIE \*néb<sup>h</sup>os ‘cloudy sky’ (p. 31)

Ved. *pāñca*, Gr. πέντε < PIE \*pénk<sup>w</sup>e ‘five’ (p. 24, 90)

Ved. *pitā*, Gr. πατήρ, Goth. *fadar* < PIE \*ph<sub>2</sub>tēr ‘father’ (p. 46)

Ved. *dhīmás*, Gr. θῦμός ‘spirit’ < PIE \*d<sup>h</sup>uh<sub>2</sub>mós ‘smoke’

Ved. *yugám*, Gr. ζυγόν < PIE \*yugóm ‘yoke’ (p. 36)

However, there are also (often disregarded) disagreements between these languages:

Gr. πῆχυς ‘forearm’ but Ved. *bāhús* ‘arm’

Gr. πέλεκυς but Ved. *paraśús* ‘axe’

Gr. πότερος, Goth. *huþar* (< PGmc \*xwáþaraz) but Ved. *katarás* ‘which of two’

Gr. τέτταρες but Ved. *catváras*, Goth. *fidwor* (< PGmc \*feðwórez) ‘four’

Though PIE accent is usually thought to have been a “pitch-accent” (like in Vedic or Ancient Greek), it is also traditionally connected with ablaut (p. 61, 68), i.e., with the

(originally) accented full grade and unaccented zero grade. Cf. the traditionally reconstructed \*dyéws ‘day sky’ (Ved. *dyáus* ‘sky’, Gr. Ζεύς) – gen. sg. \*diwós (Ved. *divás*, Gr. Διός); \*pónteHs ‘path’ (Ved. *pánthās* – gen. sg. \*pñtHós (Ved. *pathás*); \*ph<sub>2</sub>tér (p. 46) – acc. sg. \*ph<sub>2</sub>tér̄m – dat. sg. \*ph<sub>2</sub>tréy (p. 74–75); \*h<sub>1</sub>ésmi – \*h<sub>1</sub>smés (p. 93), etc. However, the question that is rarely asked is, which was first, accent or ablaut? In some cases, like \*h<sub>1</sub>smés (Ved. *smás*) or \*ph<sub>2</sub>tréy (Ved. *pitré*), it is hardly possible for the accent to be on the zero grade \*-sm- or \*-tr-. Also, although the traditional ablaut/accident scheme works well for some forms, like the athematic present (p. 93) or perfect (p. 98), it is not present in thematic presents like \*b<sup>h</sup>érete (p. 93) or *s*-stems like \*néb<sup>h</sup>esos (p. 78–79). These are often explained away as secondary but without real proof. Likewise, it is rather strange that direct correspondences such as Ved. gen. sg. *súnas* ‘dog’ and Hitt. gen. sg. *kūnaš* ‘hound-man’ (which would point to PIE gen. sg. \*kúnos from \*k<sup>w</sup>ō(n) (p. 41)) are usually interpreted as somehow secondary to Gr. κύνος,<sup>14</sup> although the Greek form is irrelevant since all Greek monosyllabic consonant stems have an automatic end-stress in the gen. sg. (unlike Vedic). The ablaut/accident scheme also has to turn to various kinds of analogies when it comes to “unexpected” stressed zero-grade forms like \*w<sup>l</sup>k<sup>w</sup>os (p. 35), \*h<sub>2</sub>ftkos (p. 35), \*(h<sub>2</sub>)w<sup>l</sup>h<sub>1/2</sub>neh<sub>2</sub> (p. 37), \*septḥ (p. 29, 90), loc. pl. \*ph<sub>2</sub>tṝsu (Ved. *pitṝsu*, Gr. dat. pl. *πατράσι* – p. 75), etc. What is more, while \*w<sup>l</sup>k<sup>w</sup>os is root-stressed, there are cases like \*deywós (p. 49) with an unstressed full-grade root. To sum up, the case for the ablaut/accident scheme does not at all look as firm as it is usually portrayed.

As noted, the traditional IE accentology holds that it is Vedic, and to a lesser extent Greek and Germanic, that is most archaic and that the PIE accent place is to be reconstructed by simple comparison of their accent positions. However, there are different ideas about IE accent. The Moscow Accentological School (p. 524) holds that it is not Vedic and Greek but Balto-Slavic (with its much more complex accentual systems, especially in derivation, and with its full-scale accentual mobility in some word types, which cannot be derived from a rather trivial “Graeco-Aryan” PIE accentual model) that is most archaic, and that PIE, like Proto-Balto-Slavic, should be reconstructed as a tone language with every syllable/morpheme either low or high.<sup>15</sup>

**Note:** I would like to thank Mislav Benić, David Mandić, Mikhail Oslon, and Petra Šoštarić for reading and commenting on the first draft of the chapter, as well as Vytautas Rinkevičius for his help with the Baltic data. Of course, all the mistakes are just mine.

## FURTHER READING

There are a number of recent general handbooks on IE linguistics, all of them containing a section on comparative phonology and differing slightly in their perspectives, though often rather short on concrete examples of sound correspondences. We shall list those in English here: Szemerényi 1996 (dated but still useful; has no laryngeals in reconstructions), Meier-Brügger 2003 (a comprehensive treatment of phonology with numerous in-text references, but the translation is at times unfortunate), Tichy 2006 (the approach of the “Erlangen School” of IE, written as a series of university lectures), Clackson 2007 (not a standard handbook but rather problem-based, discussing methodology and theoretical issues), Voyles & Barrack 2009 (an introduction focusing on the best-attested ancient IE languages), Fortson 2010<sup>2</sup> (perhaps the best English handbook and introduction, with separate chapters on all the major branches, including a discussion of historical phonology and more detail than in other general surveys), Beekes 2011<sup>2</sup> (the introduction to the “Leiden School” of IE, with glottalic theory and no PIE \*a). Gamkrelidze & Ivanov 1995,

with their glottalic theory, is comprehensive but “heterodox”. Works like Sihler 1995 or Ringe 2006, while obviously focusing on specific languages/branches, also offer information on (P)IE in general. The most comprehensive treatment of PIE segmental phonology (without accent) is still Mayrhofer’s in Cowgill & Mayrhofer 2012<sup>2</sup>. Concerning specific problems in (P)IE phonology, the number of publications (even just of more important works) is impossible to list, so we shall just mention some of the books. For laryngeal theory cf. Lindeman 1987 (though he is sometimes unorthodox) and Bammesberger 1988 (Winter 1965 is rather old but still a classic, and Voyles & Barrack 2015 is, curiously, an attempt at the refutation of the now completely established and accepted laryngeal theory), for glottalic theory Salmons 1993, and for discussion on both Vennemann 1989. Byrd’s (2015) recent book deals with the PIE syllabification but also with the PIE phonological processes and phonology in general. Collinge 1985 (with supplementary Collinge 1995 and 1999) provides an overview of sound laws and rules in PIE and IE languages, but is all too often chaotic and unclear.

## NOTES

- 1 The Austronesian language Lifu, which has *b* only in loanwords (Lenormand 1952), is often adduced as a counterexample (but cf. Gamkrelidze & Ivanov 1995: 10–11).
- 2 For Kelabit cf. Blust 1974 and 2006. Compare also a similarly unstable system of *p* : *p<sup>h</sup>* : *b<sup>h</sup>* (but no voiced stops) in the Chinese Wu language (Mallory & Adams 2006: 52).
- 3 There, one finds the opposition of *k* : *ḳ* : *k<sup>w</sup>* and *g* : *g̣* : *g<sup>w</sup>*. Cf. Эдельман 1966: 15–18 and Эдельман 1973.
- 4 Cf. also Melchert 1987 and 2012 (though his claim that Anatolian is a *centum* branch does not seem warranted).
- 5 Cf. the list in Melchert 2012: 216–217.
- 6 Cf. a typological parallel in the *q* : *k<sup>w</sup>* : *ḳ* system in the Mataco language of the Andes (Adelaar & Muysken 2004: 493).
- 7 Cf. Kortlandt 1978.
- 8 The fact that \**h*<sub>1</sub> (like \**h*<sub>2</sub>) causes aspiration in Old Indic (cf. Ved. *ásthī*, Gr. ὀστέον < PIE \**h*<sub>3</sub>*esth*<sub>1</sub> ‘bone’) would point to \**h*<sub>1</sub> having been an *h*-type sound as well (and not a glottal stop \*ʔ as some would have it).
- 9 Cf. Pinault 1982.
- 10 There are also rare quasi-ablaut non-systematic controversial correspondences like Goth. *\*tibr* ‘sacrifice, offering’ < PIE \**depr-*, compared with \**dapn-* above with \**r/n* heteroclis (p. 77–78), which is possibly a loanword, cf. Proto-Semitic \**ḏ-b-ḥ* ‘sacrifice’; PIE \**tawros* ‘bull’ (Gr. ταύρος, Lith. *taūras* ‘aurochs’, OIr. *tarb*, etc.; probably ultimately a loanword, cf. Proto-Semitic \**ḥawru(m)*) and \*(s)*tewros* (Goth. *stiur* ‘young steer’ (Eng. *steer*), Av. *staora-* ‘big cattle’ – arguably unrelated to \**tawros*), etc.
- 11 Cf. a somewhat similar analysis for a phonologically “vowelless” Modern Chinese (Pulleyblank 1984), though this only seems to prove that basically anything is possible in a “deep” and very abstract phonological analysis.
- 12 Sometimes written as \**k<sup>w</sup><sub>e</sub>twr-* in the literature.
- 13 Cf. Kapović 2006: 151–153 for more examples and details.
- 14 Cf., e.g., Kloekhorst 2008: 506.
- 15 Cf., e.g., Dybo, Nikolayev & Starostin 1978, Николаев 1989, or Дыбо 2003. A completely different, but to a point complementary (when discussing evidence for PIE as a tone language), approach to PIE as a tone language is Lubotsky 1988.

## REFERENCES

- Adelaar, Willem F. & Muysken, Pieter C. 2004, *The Languages of the Andes*, Cambridge University Press, New York
- Bammesberger, Alfred (ed.) 1988, *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*, Winter, Heidelberg
- Beekes, Robert S.P. 2011 [Revised and corrected by Michiel de Vaan], *Comparative Indo-European Linguistics*, John Benjamins Publishing Company, Amsterdam–Philadelphia
- Blust, Robert 1974, “A double counter-universal in Kelabit,” *Papers in Linguistics* 7(3/4): 309–324
- Blust, Robert 2006, “The origin of the Kelabit voiced aspirates: A historical hypothesis revisited,” *Oceanic Linguistics* 45/2: 311–338
- Byrd, Andrew Miles 2015, *The Indo-European Syllable*, Brill, Leiden–Boston
- Clackson, James 2007, *Indo-European Linguistics. An Introduction*, Cambridge University Press, Cambridge–New York
- Collinge, N.E. 1985, *The Laws of Indo-European*, John Benjamins Publishing Company, Amsterdam–Philadelphia
- Collinge, N.E. 1995, “Further laws of Indo-European,” in: Winter, Werner (ed.), *On Languages and Language*, Mouton de Gruyter, Berlin–New York: 27–52
- Collinge, N.E. 1999, “The laws of Indo-European: The state of the art (1998),” *Journal of Indo-European Studies* 27: 355–377
- Cowgill, Warren & Mayrhofer, Manfred 2012, *Indogermanische Grammatik. I. Halbband: Einleitung, II. Halbband: Lautlehre [Segmentale Phonologie des Indogermanischen]*, Universitätsverlag Winter GmbH, Heidelberg
- Дыбо, Владимир А. 2003, “Балто-славянская акцентологическая реконструкция и индоевропейская акцентология,” *Славянское языкознание. XIII международный съезд славистов. Любляна, 2003 г. Доклады российской делегации*, Москва: 131–161
- Dybo, V.A., Nikolayev, S.L. & Starostin, S.A. 1978, “A tonological hypothesis on the origin of paradigmatic accent systems,” *Estonian Papers in Phonetics* 1978: 16–20
- Эдельман, Джой Иосифовна 1966, *Язулямский язык*, Издательство “Наука,” Москва
- Эдельман, Джой Иосифовна 1973, “К типологии индоевропейских гуттуральных,” *Известия Академии наук СССР. Серия литературы и языка* XXXII/6: 540–546
- Fortson, Benjamin W. IV 2010<sup>2</sup>, *Indo-European Language and Culture. An Introduction*, Wiley-Blackwell, Malden–Oxford
- Gamkrelidze, Thomas V. & Ivanov, Vjačeslav V. 1995, *Indo-European and the Indo-Europeans. A Reconstruction and Historical Analysis of a Proto-Language and a Proto-Culture*, Mouton de Gruyter, Berlin–New York [translation of 1984]
- Kapović, Mate 2006, *Reconstruction of Personal Pronouns in Balto-Slavic with Emphasis on Accentuation*, University of Zadar, Zadar [PhD dissertation]
- Kloekhorst, Alwin 2008, *Etymological Dictionary of the Hittite Inherited Lexicon*, Brill, Leiden–Boston
- Kortlandt, Frederik 1978, “I.-E. palatovelars before resonants in Balto-Slavic,” in: Fisiak, Jacek (ed.), *Recent Developments in Historical Phonology*, Mouton, The Hague: 237–243
- Kortlandt, Frederik 1985, “Proto-Indo-European glottalic stops: The comparative evidence,” *Folia Linguistica Historica* 6/2: 183–201 [reprinted in: Kortlandt, Frederik 2010, *Studies in Germanic, Indo-European and Indo-Uralic*, Rodopi, Amsterdam–New York: 53–66]

- Lehmann, Winfred P. 1993, *Theoretical Bases of Indo-European Linguistics*, Routledge, London–New York
- Lenormand, Maurice H. 1952, “The phonemes of Lifu, Loyalty islands: The shaping of a pattern,” *Word* 8/3: 252–258
- Lindeman, Fredrik Otto 1987 [1997], *Introduction to the Laryngeal Theory*, Norwegian University Press, Oslo
- LIV 2001<sup>2</sup>, *Lexikon der Indogermanischen Verben*, Helmut Rix (ed.), Dr. Ludwig Reichert Verlag, Wiesbaden
- Lubotsky, Alexander M. 1988, *The System of Nominal Accentuation in Sanskrit and Proto-Indo-European*, Brill, Leiden, New York
- Lubotsky, Alexander M. 1989, “Against a Proto-Indo-European phoneme \*a,” in: Vennemann, Theo (ed.), Leiden: de Gruyter: 53–67
- Николаев, Сергей Л. 1989, “Балто-славянская акцентуационная система и ее индоевропейские истоки,” *Историческая акцентология и сравнительно-исторический метод*, Москва: 46–109
- Mallory, James P. & Adams, Douglas Q. 2006, *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*, Oxford University Press, Oxford–New York
- Meier-Brügger, Michael 2003 [With contributions by Matthias Fritz and Manfred Mayrhofer], *Indo-European Linguistics*, Walter de Gruyter, Berlin–New York [translation of the 2002 German edition]
- Melchert, Craig H. 1987, “PIE velars in Luvian,” in: Watkins, Calvert (ed.), *Studies in Memory of Warren Cowgill (1929–1985). Papers from the Fourth East Coast Indo-European Conference (Cornell University, June 6–9, 1985)*, Walter de Gruyter, Berlin–New York: 182–204
- Melchert, Craig H. 2012, “Luvo-Lycian Dorsal stops revisited,” in: Sukač, Roman & Šefčík, Ondřej (eds.), *The Sound of Indo-European 2. Papers on Indo-European Phonetics, Phonemics and Morphophonemics*, Lincom Europa, München: 206–218
- Pinault, Georges-Jean 1982, “A neglected phonetic law: The reduction of the Indo European laryngeals in internal syllables before yod,” in: Ahlqvist, Anders (ed.), *Papers from the 5th International Conference on Historical Linguistics*, John Benjamins, Amsterdam: 265–272
- Pulleyblank, Edwin G. 1984, “Vowelless Chinese? An application of the three-tiered theory of syllable structure,” in: *Proceedings of the Sixteenth International Conference on Sino-Tibetan Languages and Linguistics*, vol. ii, University of Washington, Seattle: 568–619
- Ringe, Don 2006, *From Proto-Indo-European to Proto-Germanic*, Oxford University Press, Oxford–New York
- Salmons, Joseph C. 1993, *The Glottalic Theory: Survey and Synthesis*, Institute for the Study of Man, McLean
- Sihler, Andrew 1995, *New Comparative Grammar of Greek and Latin*, Oxford University Press, New York–Oxford
- Szemerényi, Oswald J.L. 1996, *Introduction to Indo-European Linguistics*, Clarendon Press–Oxford, Oxford–New York [translation of the 1990 German edition]
- Tichy, Eva 2006, *A Survey of Proto-Indo-European*, Hempen Verlag, Bremen
- Vennemann, Theo (ed.) 1989, *The New Sound of Proto-Indo-European : Essays in Phonological Reconstruction*, Mouton De Gruyter, Berlin–New York
- Voyles, Joseph & Barrack, Charles 2009, *An Introduction to Proto-Indo-European and the Early Indo-European Languages*, Slavica Publishers, Bloomington



Voyles, Joseph & Barrack, Charles 2015, *On Laryngealism. A Coursebook in the History of a Science*, Lincom GmbH, München  
Winter, Werner (ed.) 1965, *Evidence for Laryngeals*, Mouton, The Hague

### **SYMBOLS USED IN THIS AND THE FOLLOWING CHAPTER**

- C – any consonant
- V – any vowel
- R – any resonant
- Ṛ – any syllabic resonant
- Ḳ – any velar
- <> – orthographic writing
- // – phonological writing
- [] – phonetic writing
- # – beginning/end of a word
- \* reconstructed
- + not directly attested (in a text, etc.) but inferred
- < (is) derived from
- > (has) become
- ⇐ (is) a loanword from; non-phonetic development
- ⇒ (is) a loanword into
- ~ connected with, related to
- lit. – literally

# PROTO-INDO-EUROPEAN MORPHOLOGY

*Mate Kapović*

## INTRODUCTION

Proto-Indo-European, as well as all early attested Indo-European languages, had a rich and complex synthetic morphology. This morphology was often simplified in later languages – e.g., no modern IE language has all the eight cases PIE had, and some have none. However, new morphological forms also appeared – e.g., most modern IE languages have a separate future tense (though not always synthetic), which did not exist in PIE. Forms and endings are usually easier to reconstruct than their functions and exact meaning. The focus of this short overview will be on the reconstruction of PIE forms and paradigms and not on the exact (and sometimes disputable) function and meaning of those forms. Obviously, not all reflexes and developments in all descendant languages can be dealt with here. As a rule, only the most archaic attested forms (even if rare) are adduced, with the point of reconstructing PIE forms and paradigms. The focus is more on PIE itself, rather than on the development and situation in separate IE branches (dealt with in separate chapters of this volume). Non-inflected parts of speech, though not strictly part of morphology, are also dealt with in the end, while word formation will be tackled only marginally throughout the chapter.

## ABLAUT

Ablaut is a regular morphonological alternation of vowels in PIE morphemes. Depending on its form (either as part of a morphological paradigm or derivational process), a PIE root, suffix (rarely a prefix), or ending could have had different shapes. The following vowels were in alternations: the full grade (\*e, \*o), the lengthened or long grade (\*ē, \*ō), and the zero grade (\*Ø). In the case of \*eR in the full grade (\*R being any resonant/glide), the zero grade was \*R/Ṛ (i.e., \*ey – \*oy – \*ēy – \*ōy – \*y/i; \*em – \*om – \*ēm – \*ōm – \*m/ṃ, etc.), depending on syllable structure (p. 31). Thus, we have PIE \*h<sub>1</sub>es-ti ‘it is’ but \*h<sub>1</sub>s-enti ‘they are’ (with \*h<sub>1</sub>es-/h<sub>1</sub>s- alternation). The rules determining the appearance of a certain ablaut form were quite complex in PIE. Some ablaut alternations are seen even in modern IE languages (e.g., *sing* – *sang* – *sung* in English), but they are best preserved in early attested languages. Cf. all five grades in the suffix in these Greek forms:

πά-τερ! ‘father!’ (\*e-grade)

ἄ-πα-τορ! ‘orphan (fatherless)!’ (\*o-grade)

πα-τρ-ός ‘of the father’ (zero grade)

πα-τήρ ‘father’ (\*ē-grade)

ἄ-πά-τωρ ‘orphan’ (\*ō-grade)

In certain roots, resonants can change their place – cf. \*dyew-/deyw- in \*dyēws (p. 40, 56) and \*deywos (p. 49) – this is called *Schwebeablaut* ‘floating ablaut’ (cf. Anttila 1969). PIE verbal, nominal, and verbal/nominal roots had a CVC structure (with possible consonant clusters – p. 52–54), in most of which the ablaut scheme in the -V- part is attested.

## NOUNS

Late (post-Anatolian) PIE nominals had three genders – masculine, feminine, and neuter – preserved in most early IE languages. The feminine gender is possibly a common PIE innovation (which is controversial) after the early split of the Anatolian branch, since Anatolian exhibits only a distinction of animate/common and inanimate/neuter nouns (p. 178). PIE had three numbers – singular, plural, and dual (the latter often considered a late addition to the PIE number system). Adjectives and pronouns agreed in number and gender with nouns, and verbs agreed with them in number (but cf. the curious case of the neuter nominative “plural” – p. 65, 166). Morphologically, adjectives behaved exactly like nouns as far as declension goes, and they can be treated together, while pronouns had some special case endings and traits. There were eight cases – nominative, genitive, dative, accusative, vocative, locative, instrumental, and ablative.

PIE nouns had the following formation: (prefix +) root (+ suffix) (+ suffix) + ending. A noun had to have a root and an ending (even if it was a zero morpheme). Prefixes (like \*n̥- ‘non-’), which were rare (p. 156–157), occurred in some prefix derivatives (p. 79–80) – these can be interpreted as the first part of a compound, in which case PIE would have had no prefixes. Words could have more than one suffix – cf. \*pōd-s ‘foot’ (root + ending), \*pot-i-s ‘lord’ (root + suffix + ending), and \*suH-n-u-s ‘son’ (root + suffix + suffix + ending). There were also some nouns with reduplication, e.g., \*k<sup>w</sup>e-k<sup>w</sup>l-o-s ‘wheel’.

Many nouns have special declension suffixes directly preceding the case endings. Thus, we speak of different nominal stems: *o*-stems (\*h<sub>1</sub>ekw-o-s ‘horse’); *eh<sub>2</sub>*-stems (\*(h<sub>2</sub>)w|<sub>h</sub>1n-eh<sub>2</sub>-Ø ‘wool’); root nouns (\*pōd-s, no suffix); *i*-stems (\*h<sub>3</sub>ew-i-s ‘sheep’); *u*-stems (\*suHn-u-s); resonant stems – *r*-stems (\*ph<sub>2</sub>t-ēr-Ø ‘father’), *n*-stems (\*h<sub>2</sub>ekm-ō(n)-Ø ‘stone’), *m*-stems (\*d<sup>h</sup>eg<sup>h</sup>-ō(m)-Ø ‘earth’), and *l*-stems (\*h<sub>2</sub>eb-ōl-Ø ‘apple’); laryngeal stems (\*pont-eH-s ‘path’); plosive stems (\*(h<sub>2</sub>)nep-ōt-s ‘nephew’); and *s*-stems (\*neb<sup>h</sup>-os-Ø ‘cloudy sky’). These declension suffixes (\*suHn-u-s) can be distinguished from the strictly derivational suffixes (\*suH-n-us, cf. the verbal root \*suH- ‘give birth’ > Ved. *sūte* ‘gives birth’) that form the extended root. It goes without saying that sometimes it is impossible to ascertain the exact morphological/derivational shape of a word; e.g., is \*ponteHs to be analyzed as \*pon-teH-s or \*pont-eH-s? The declension suffixes are not necessarily always the same, cf. the different ablaut shapes in suffixal \*-ē(n), \*-ō(n), \*-n̥ in the *n*-stems (p. 77).

*O*- and *eh<sub>2</sub>*-stems are called thematic stems (after the so-called thematic vowel \*-e/o-), and the other stems athematic. All athematic stems can be called consonantal stems (since \*i/u can be considered syllabic glides – p. 34). In thematic stems, there was no ablaut alternation in the root and endings – they had the same shape in all the cases (the only ablaut alternation in thematic stems was the very limited qualitative \*-e-/o- alternation of the thematic vowel in *o*-stems, not unlike athematic *s*-stems). For instance, the root in \*h<sub>1</sub>ekwos was in the *e*-grade in all the cases (gen. sg. \*h<sub>1</sub>ekwosyo, dat. sg. \*h<sub>1</sub>ekwōy, etc.), and the root in \*w<sub>l</sub>k<sup>w</sup>os ‘wolf’ was always in the zero grade (gen. sg. \*w<sub>l</sub>k<sup>w</sup>osyo, dat. sg. \*w<sub>l</sub>k<sup>w</sup>ōy, etc.). In athematic stems, all morphemes (root, suffix, ending) could potentially (according to their ablaut type) change their ablaut shape (i.e., they could have

one ablaut of the root in one case and another one in another case), though this did not occur in all forms and types (the *s*-stems were the most atypical). For instance, the gen. sg. of \*pōnt-eH-s was \*pōnt-H-e/os with a different grade in both the root and (the pre-sumed) suffix. The same opposition of thematic/athematic occurs in verbs as well (p. 93).

Case endings differed slightly depending on the type of declension and gender (*o*-stems differed the most from all other stems, and their specific endings are listed separately below):

	<b>singular</b>	<b>plural</b>	<b>dual</b>
nom.	*-s (m./f.), *-Ø (n.)	*-es (m./f.), *-h <sub>2</sub> (n.), *-Ø (n.)	*-h <sub>1</sub> (e) (m./f.), *-ih <sub>1</sub> (n.)
voc.	*-Ø	= nom.	= nom.
acc.	*-m (m./f.), *-Ø (n.)	*-ns (m./f.), = nom. (n.)	= nom.
gen.	*-(e/o)s	*-om	(?) *-h <sub>1</sub> s, (?) *-u
abl.	= gen.	= dat.	= dat.
dat.	*-ey	*-b <sup>h</sup> os/*-mos	(?) *-b <sup>h</sup> (y)oh <sub>1</sub> /*-moh <sub>1</sub>
loc.	*-i, *-Ø	*-su	(?) *-u
instr.	*-(e)h <sub>1</sub>	*-b <sup>h</sup> i(s)	= dat.

*O*-stems differed in these endings: neuter nom./acc. sg. \*-m; masculine gen. sg. \*-syo; abl. sg. \*-ot; masculine loc. pl. \*-isu; masculine loc. pl. \*-eys. For athematic endings cf. p. 69–70.

In neuter nouns, nom./acc./voc. were always the same in all numbers. In the dual, only nom./acc./voc. can be reconstructed with certainty. The ablative had a special ending only in the sg. of *o*-stems; elsewhere, it was identical to the gen. (sg.) or dat. (pl./du.). In some cases, it is possible to speculate on the origin of the endings. The dat. sg. \*-ey and loc. sg. \*-i look like ablaut variants (and perhaps were just that originally in pre-PIE) but are distinct and invariant in the last stage of PIE (ablaut variants are found only in athematic gen./instr. sg. – p. 68). Acc. pl. \*-ns surely derives from agglutinative \*-m-s (acc. sg. + nom. pl.), and some reconstruct \*-ms as the ending,<sup>1</sup> while the endings of the dat./loc./instr. pl. and dat./abl./instr. du. are of clear postpositional origin. In Celtic, Italic, Indo-Iranian, and Greek one finds endings with \*-b<sup>h</sup>- in the dat./instr. pl./du. (with different variant forms), while Balto-Slavic and Germanic have \*-m- here (the exact relation of these is unclear).<sup>2</sup> The reconstruction here is rather provisional. Indo-Iranian points to the instr. pl. \*-b<sup>h</sup>i-s (parallel to the BSl. instr. pl. \*-mi-H-s), while Greek (Mycenaean and Homeric) has the instr. pl. (used also in other functions in Greek) <-pi>/-φι without the plural \*-s (p. 302) – cf. also the Arm. instr. sg. ending -v < \*-b<sup>h</sup>i (p. 434–435) and OCS instr. sg. -mb < \*-mi (with \*-m-). This could be originally related to the Germanic preposition \*bi (> Eng. *by*), originally a postposition (case endings often appear out of former postpositions). Indo-Iranian points to proto-variants \*-b<sup>h</sup>y-os for the dat./abl. pl. and \*-b<sup>h</sup>y-oh<sub>1</sub> for the dat./abl./instr. du., with \*-b<sup>h</sup>i- where other languages have \*-b<sup>h</sup>- (or \*-m-). Different accounts are possible here (e.g., an original dat. \*-mos but instr. pl. \*-b<sup>h</sup>i(s), etc.) – provisionally, we reconstruct the dat. pl. as \*-b<sup>h</sup>os/-mos and the dat. du. as \*-b<sup>h</sup>(y)oh<sub>1</sub>/-moh<sub>1</sub>. The nom. du. perhaps has its origin in the instr. sg. (Schmalstieg 1998), with an ellipsis of the type \*w<sub>1</sub>k<sup>w</sup>os w<sub>1</sub>k<sup>w</sup>oh<sub>1</sub> ‘wolf with a wolf’ > (\*w<sub>1</sub>k<sup>w</sup>os) w<sub>1</sub>k<sup>w</sup>oh<sub>1</sub> ‘two wolves’ – cf. the Vedic type *Mitrā* (nom. du.) (. . .) *Váruṇo* (nom. sg.) ‘Mitra and Varuna’ and the ellipsis in Ved. *dyāvā* ‘heaven and earth’ (literary: ‘two heavens’) and the like cases in Greek and Slavic. The \*-h<sub>1</sub> in the dual forms \*-ih<sub>1</sub> and \*-b<sup>h</sup>y/moh<sub>1</sub> is assumed by analogy to the nom. du. \*-h<sub>1</sub>. In the gen./loc. du. the \*-o-w from *o*-stems seems to have spread to other stems (otherwise, \*-u should be reconstructed as an *o*-stem variant, and

\*-ow for other stems). In this chapter, we will not reconstruct all dual forms everywhere (especially when there are no reflexes of the original forms).

### Symbols (used in tables and lists of forms)

{form or ending not related to the original PIE one or completely transformed}, [form or ending reflecting the PIE one but with a different function or taken from another form in the paradigm], \an irregular phonological change, analogy, or a non-transparent addition in the form or ending\, |a secondary addition to a form or ending|

### Thematic stems

#### *O-stems*

*O*-stems were mostly masculine (like \*w<sup>h</sup>k<sup>w</sup>os) or neuter (like \*yugom ‘yoke’). There were also some feminine *o*-stems like \*snusos ‘daughter-in-law’ (p. 36). Masculine and neuter nouns had different endings only in the nom./voc. sg. and nom./acc./voc. pl./du. All cases had thematic \*-o- except the voc. sg., which had \*-e and a zero ending.

TABLE 1.10 IE *O*-STEMS (SINGULAR)

sg.	PIE	Ved.	Gr.	OLat.	OCS	Lith.	Goth.
nom.	*-o-s	-as <sup>a</sup>	-oς	-os	-b	-as	-s
voc.	*-e	-a	-ε	-e	-e	-e	-Ø
acc.	*-o-m	-am	-ov	-om	-b	-q	-Ø
gen.	*-o-syo	-asya	Hom. -oio	-osio		<sup>g</sup>	{-is}
abl.	*-o-ot	-ād	Dor. ὀπῶ <sup>c</sup>	-ōd	[gen. -a]	[gen. -o]	
dat.	*-o-ey	-āy a  <sup>b</sup>	-φ	-oi	-u	-ui	-a
loc.	*-o-y	-e	οἴκοι <sup>d</sup>	domī <sup>e</sup>	-ē	namiē <sup>d</sup>	
instr.	*-o-h <sub>1</sub>	RV -ā			{-omb} <sup>f</sup>	-ù	OHG -u

<sup>a</sup> The final -s in Old Indic (in this and other forms) is often written as -h (a different sandhi-variant); <sup>b</sup> cf. GAv. -āi; <sup>c</sup> ‘whence’ (adv.); <sup>d</sup> ‘at home’ (adv.) (p. 302); <sup>e</sup> Class. Lat. ‘at home’ (synchronously gen.); <sup>f</sup> cf. adv. *vbčēr-a* ‘yesterday’; <sup>g</sup> OPruss. -as (< \*-os(y)o?).

The gen. sg. ending \*-osyo (cf. also Lep. -oiso, Arm. -oy, HLuw. -as(s)i), different from the usual \*(e/o)s found in other declensions, is usually compared to the pronominal ending \*-eso (p. 81), though the exact connection is not clear. The Latin and Old Indic abl. ending (also Clb. -uō) can be derived from both \*-t and \*-d (p. 53), but the possible connection to Hitt. abl. -az < \*-ot-i, Toch. A -s < \*-ti (p. 462), and OCS preposition *otb* ‘of’ would point to the original PIE abl. \*-ot (though \*-od is often reconstructed for PIE).

TABLE 1.11 IE *O*-STEMS (PLURAL)

pl.	PIE	Ved.	Gr.	OLat.	OCS	Lith.	Goth.
nom.	*-o-es	-ās	{-oi}	{-oe}	{-i}	{-ai}	-os
acc.	*-o-ns	-ān <sup>b</sup>	dial. -ovς	-ōs	-y/-e	-ūs	-ans

pl.	PIE	Ved.	Gr.	OLat.	OCS	Lith.	Goth.
gen.	*-o-om	RV <i>-ām</i> <sup>c</sup>	-ov	<i>-om</i>	-o <sup>d</sup>	-ū	{-e}
dat.	*-o-b <sup>h</sup> os <sup>a</sup>	-e <sup>h</sup> b <sup>h</sup> lyas			-om <sup>o</sup>	OLith. <i>-am\us</i>	<i>-am</i>
loc.	*-o-ysu	<i>-eṣu</i>	dial. -OI Σ \		-ěx <sup>o</sup>	dial. <i>-iesu</i> <sup>e</sup>	
instr.	*-o-ey-s	<i>-ais</i>	[dat. -oiç]	[dat. <i>-ois</i> ]	-y	<i>-ais</i>	

<sup>a</sup> BSL/Gmc variant \*-mos, Ilr. \*-b<sup>h</sup>yos; <sup>b</sup> sandhi-variant: *-āms*, cf. GAv. sandhi-form *-as*; <sup>c</sup> only four times in the fossilized collocation *dev-āñ jānman-* ‘race of gods’; <sup>d</sup> PSI. \*-ō; <sup>e</sup> (cf. p. 502).

Two adjacent vowels contracted: dat. sg. (\*-ōy < \*-o-ey), abl. sg. (\*-ōt < \*-o-ot), nom. pl. (\*-ōs < \*-o-es), instr. pl. (\*-ōys < \*-o-ey-s), and gen. pl. (\*-ōm < \*-o-om). It is possible that these endings were still (at least as variants or dialectally) uncontracted in PIE, judging by occasional disyllabic scansion in Indo-Iranian, and perhaps by the development of Lith. gen. *-o* < \*-ā (p. 41) < BSL. \*-a-at < PIE abl. \*-o-ot (with contraction after PIE, \*o > BSL. \*a?). In many languages/branches, the original ending \*-ōs (cf. also Osc. *núvlan-ús* ‘inhabitants of Nola’) was replaced by the pronominal \*-oy (\*toy w<sub>l</sub>k<sup>w</sup>ōs ‘those wolves’ ⇒ \*toy w<sub>l</sub>k<sup>w</sup>oy, p. 81). The \*-b<sup>h</sup>-/-m- in the dat. (and elsewhere) is problematic (p. 63), but \*-os is clear from BSL. \*-mos (OCS *-m<sup>o</sup>*, OLatv. *-ms*), Lat. *-bus* (in athematic stems), Celt. \*-b<sup>h</sup>os (Cib. *-bos*), and Ven. *-bos* (*louderobos* ‘to the children’), supported by Ilr. \*-b<sup>h</sup>-y-os. The loc. \*-i-su probably has the additional \*-i- from the singular (cf. the simple ending \*-su in other declensions), and \*-ey-s in the instr. pl. can tentatively be connected to the Slavic conjunction *i* ‘and’.

TABLE 1.12 IE O-STEMS (NEUTER)

neuter	PIE	Ved.	Gr.	OLat.	OCS	Goth.	Hitt.
nom./acc./voc. sg.	*-o-m	<i>-ám</i>	-óv	<i>-om</i>	-o\	-Ø	<i>-an</i>
nom./acc./voc. pl.	*-e-h <sub>2</sub>	<i>-ā</i>	-á\	-a\	-a	-a	<i>-a</i>

The nom./acc. forms are always the same in neuter nouns. In *o*-stems, the neuter nom./acc. sg. has the same form as the masculine acc. sg. (\*-om). This, and the fact that nom./acc. are always the same in neuter nouns, undoubtedly derives from the pragmatic use of neuter (inanimate) nouns, which are rarely the subjects of transitive verbs. For this reason (and others), pre-PIE is considered by some to have been an ergative language (cf. 162–164). The ending \*-eh<sub>2</sub> of the nom./acc./voc. pl. is identical to the nom. sg. of feminine \*eh<sub>2</sub>-stems. That is due to the fact that the “plural” of neuter nouns is originally a collective, not a plural *per se*. That is why in IE languages one finds the \*-eh<sub>2</sub>- plural forms for masculine nouns as well, cf. Av. regular masculine nom. pl. in *-a* (p. 275) or Lat. nouns like *iocus* ‘joke’ – pl. both *iocī* and *ioca* (similar cases exist in Iranian, Greek, Anatolian, Slavic, etc.). In Greek, Anatolian, and Gatha-Avestan, the plural of neuter nouns agrees with verbs in the 3 sg. This is known as the τὰ ζῷα τρέχει ‘the animals (pl.) run (sg.)’ rule in Greek, and is an archaic PIE feature stemming from the said fact that later neuter nominative plurals were originally collective singulars.

TABLE 1.13 IE *O*-STEMS (DUAL)

du.	PIE	Ved.	GAv.	Gr.	OCS	Lith.
nom./acc./voc. (m.)	*-o-h <sub>1</sub>	-ā	-ā	-ω	-a	-ù
nom./acc./voc. (n.)	*-o-yh <sub>1</sub>	-e		{-ω}	-ě	
gen.	(?) *-o-h <sub>1</sub> s	[-ayos]	- aii ā	{-ouv} <sup>b</sup>	[-u]	
loc.	*-o-w	[-ay o s] <sup>b</sup>	- aii ō <sup>b</sup>		-u	(?) <i>pusiáu</i> <sup>c</sup>
dat./abl./instr.	(?) *-o-b <sup>b</sup> (y)oh <sub>1</sub> <sup>a</sup>	-ā bhyā m	-ā biīā	{-ouv} <sup>b</sup>	-oma	-ām/-aĩ

<sup>a</sup> BSL. \*-omoh<sub>1</sub>; <sup>b</sup> cf. pronominal forms – p. 86; <sup>c</sup> dial. ‘in two/half’.

Old Church Slavic and Vedic point to \*-ow in both gen. and loc. (the gen. \*-s in Vedic is probably secondary), but Avestan has gen. -ā < \*-ās (presumably from \*-o-h<sub>1</sub>-s with nom. \*-h<sub>1</sub>- + gen. \*-s) and loc. -ō < \*-aw, which is possibly archaic (Beekes 2011: 217). The loc. (and gen.) ending was originally perhaps \*-u, thus \*-o-u in *o*-stems, but this \*-ow was later, possibly already in PIE, generalized in other stems as well.

### *Eh<sub>2</sub>-stems*

*Eh<sub>2</sub>-stems* (which were always feminine) seem to be a post-Anatolian innovation, since they do not appear there. In late PIE, these stems are correlated to *o*-stems in adjectives: \*newos (m./newom (n.) and \*neweh<sub>2</sub> (f.) ‘new’. The post-Anatolian emergence of the new feminine gender could have originated from the identification of the collective \*-eh<sub>2</sub> (and possessive \*-ih<sub>2</sub>, cf. Lat. *o*-stem gen. sg. -ī) suffix with the accidental \*-h<sub>2</sub> in the originally athematic word for ‘woman’: \*g<sup>w</sup>enh<sub>2</sub> (Ved. *jāni*, OIr. *bé*, p. 77, 369).<sup>4</sup>

TABLE 1.14 IE *EH<sub>2</sub>-STEMS (SINGULAR)*

sg.	PIE	Ved.	Gr.	Lat.	OCS	Lith.	Goth.
nom.	*-eh <sub>2</sub>	-ā	-ā	-ā\	-a	-à	-a
voc.	*-e(h <sub>2</sub> )	{-e}	-ā <sup>a</sup>	= nom.	-o	-a	= nom.
acc.	*-eh <sub>2</sub> -m	-ām	-āv	-am	-ǫ	-q	-a
gen.	*-eh <sub>2</sub> -es	- āy ās	-āς	OLat. -ās	[-y/ǫ]	-ðs	-os
dat.	*-eh <sub>2</sub> -ey	- āy ai	-āi	-ae	-ě	-ai	-ai
loc.	*-eh <sub>2</sub> -i	{-āyām}		-ae	-ě	[-ai- p ] <sup>c</sup>	
instr.	*-eh <sub>2</sub> -eh <sub>1</sub>	-ā	κρυφᾶ <sup>b</sup>		{-oǫ}	-ā\	

<sup>a</sup> Hom. νόμφ-ᾶ!; <sup>b</sup> adv. ‘in secret’; <sup>c</sup> with a postposition -p in the new adessive.

All languages except Latin point to \*-ā (cf. also Umbr./Osc. -o < \*-ā and Celtic – p. 369) in the nom., which is then interpreted as older \*-e-h<sub>2</sub> (p. 43), though the laryngeal is nowhere attested (\*-e- is a thematic vowel). However, the existence of other feminine endings (\*-i-h<sub>2</sub>-(s) and \*-u-h<sub>2</sub>-s, p. 67, 74) indicates a laryngeal. Other cases have this \*-eh<sub>2</sub>- (which has a zero ending in the nom.) + general endings: acc. \*-eh<sub>2</sub>m (not \*\*-eh<sub>2</sub>m!) > \*-ām (Stang’s Law – p. 70); gen. \*-eh<sub>2</sub>es > \*-ah<sub>2</sub>as > \*-aas (p. 43) > \*-ās; dat. \*-eh<sub>2</sub>ey > \*-ah<sub>2</sub>ay > \*-aay > \*-āy; loc. \*-eh<sub>2</sub>i > \*-ah<sub>2</sub>i > \*-ay; instr. \*-eh<sub>2</sub>eh<sub>1</sub> > \*-ah<sub>2</sub>ah<sub>1</sub> > \*-aā > \*-ā, etc. Some endings are formally reconstructed on a structural basis in their laryngealistic and pre-contraction shape. Greek and Balto-Slavic (cf. also Umbr. voc.

*tursa!* (name of a goddess)) point to \*-ǎ in the voc. In laryngeal terms, this is usually explained as \*-e(h<sub>2</sub>), with the laryngeal coloring the preceding vowel (\*-a(h<sub>2</sub>)) but then dropping before a pause without lengthening (cf. also Ved. *devi!* < \*-i(h<sub>2</sub>) from *devī* ‘goddess’ < \*-ih<sub>2</sub>, see below).

TABLE 1.15 IE EH<sub>2</sub>-STEMS (PLURAL)

pl.	PIE	Ved.	Gr.	Lat.	OCS	Lith.	Goth.
nom.	*-eh <sub>2</sub> -es	-ās	{-aī}	OLat. -as	[-y/-e]	-os	-os
acc.	*-eh <sub>2</sub> -(n)s	-ās	dial. -ávς	-ās	-y/-e	-ās	-os
gen.	*-eh <sub>2</sub> -om	-[ān]ām	dial. -ávς	{-ārum} <sup>c</sup>	-o <sup>d</sup>	-ū	-o
dat.	*-eh <sub>2</sub> -b <sup>h</sup> os <sup>a</sup>	-ābhyas	{-aīς}	{-īs}	-amъ	OLith. -om\us	-om
loc.	*-eh <sub>2</sub> -su	-āsū			-aʎʎ	dial. -osu	
instr.	*-eh <sub>2</sub> -b <sup>h</sup> i(s) <sup>b</sup>	-ābhis	Myc. -āp <sup>h</sup> i		-ami	-omis	

<sup>a</sup> BSL./Gmc \*-mos, Iir. \*-b<sup>h</sup>yos; <sup>b</sup> Gr. \*-b<sup>h</sup>i, Iir. \*-b<sup>h</sup>is, BSL. \*-miHs; <sup>c</sup> cf. pronominal \*-som (p. 81); <sup>d</sup> PSlav. \*-ṛ.

Some IE languages (Vedic, Gothic) show no trace of the original \*-n- in the acc. pl., so it is possible to assume that this was already a PIE change (at least in some dialects). The gen. ending \*-eh<sub>2</sub>-om is formally reconstructed, since, if we disregard the secondary forms, all we have are a few reflexes pointing to a contracted long \*-ōm (Goth. -o, Lith. -ū, PSl. \*-ṛ).

TABLE 1.16 IE EH<sub>2</sub>-STEMS (DUAL)

du.	PIE	Ved.	OCS	Lith.
nom./acc./voc.	*-eh <sub>2</sub> -ih <sub>1</sub>	-e	-ě	-i <sup>c</sup>
gen./loc.	*-eh <sub>2</sub> -u <sup>a</sup>	{-ayos}	-u	
dat./abl./instr.	(?)*-eh <sub>2</sub> -b <sup>h</sup> yoh <sub>1</sub> <sup>b</sup>	-ābhyā m	-ama	-ómi/-ōm

<sup>a</sup> Or analogical \*-eh<sub>2</sub>-ow (p. 63); <sup>b</sup> BSL. \*-eh<sub>2</sub>moh<sub>1</sub>; <sup>c</sup> cf. -ie-ji-dvi in definite adjectives with -ie- < \*-ayH-

The nom./acc. endings point to \*-ay (< \*-aī < \*-a(h<sub>2</sub>)ih<sub>1</sub>). The eh<sub>2</sub>-stems have the dual \*-ih<sub>1</sub> (like the neuter – perhaps due to the connection of the feminine sg. and neuter pl., p. 65), not \*-h<sub>1</sub>(e) like other feminine stems.

Two feminine \*-ih<sub>2</sub>-types can also be reconstructed (PIE \*deyw-ih<sub>2</sub> ‘goddess’ > Ved. *devī*; cf. Lith. *patī* ‘wife’; PIE \*w<sub>1</sub>lk<sup>w</sup>-ih<sub>2</sub>-s ‘she-wolf’ > Ved. *vṛkīś*, p. 236–237); one of the types had the apophonic alternation \*-ih<sub>2</sub>-/-yeh<sub>2</sub>- in the suffix (Ved. dat. sg. *devyāi* < \*deyw-yeh<sub>2</sub>-ey but *vṛkīe* < \*w<sub>1</sub>lk<sup>w</sup>-ih<sub>2</sub>-ey).

### Athematic stems

#### PIE athematic ablaut declension types

As already said, in athematic nouns all morphemes (roots, suffixes, endings) could exhibit ablaut alternations – for instance, the suffix was in the lengthened grade in \*ph<sub>2</sub>tēr ‘father’, in the zero grade in \*suHnus ‘son’ and the gen. sg. \*ph<sub>2</sub>tre/os, in the full grade



in the gen. sg. \*suHnows, while the ending was in the full grade in the gen. sg. \*ph<sub>2</sub>tré/ós but in the zero grade in the gen. sg. \*suHnows. PIE had quite a number of complex ablaut declension types (ADTs), which are difficult to reconstruct.

In the case of root alternations, the original ADTs can sometimes, but very rarely, be reconstructed from direct reflexes in some language (e.g., \*ponteHs – gen. sg. \*pntHos ‘path’ can be reconstructed from Ved. *pánthās* – *pathás*). Usually, they have to be reconstructed from different reflexes in various languages, where one of the alternate PIE ablaut grades had usually been generalized – e.g., if one compares Lat. *aurōra* (< \*ausōs-) with Ved. *uṣás* (gen. sg. *uṣás*), we can conclude that PIE had ablaut variants \*h<sub>2</sub>ewsōs-, \*h<sub>2</sub>usōs-, and \*h<sub>2</sub>us(s)- (by additional analysis in the context of the whole PIE declension system, we come to the conclusion that the original pattern was probably nom. sg. \*h<sub>2</sub>ewsōs – gen. sg. \*h<sub>2</sub>use/ós ‘dawn’, p. 51). However, already in PIE, as far as can be reconstructed, many stem types (e.g., *i*-stems, most *r*-stems, etc.) did not exhibit ablaut alternations in the root (which may or may not be an innovation in comparison to the pre-PIE period). The alternations in suffixes (cf. Lat. *pater* ‘father’ – gen. sg. *patris*), more common in PIE, were also more likely to be preserved, though here one of the types usually tends to be generalized in the daughter languages.

The standard theory of ADTs connects the ablaut types with the supposed PIE accentual alternations, the usual supposition being that the full (especially \*e) grade goes with accentedness, while the zero grade should be unaccented. However, there are many problems with this standard account, as well as with the reconstruction of PIE accent in general (p. 55–56), and its relation to ablaut patterns. The problem with the standard theory is that it usually (at least unconsciously) tries to reconstruct the ablaut types not of the latest stage of PIE, but rather of some “original” supposed pre-PIE stage, and thus often operates with the supposed, but completely speculative, forms, like \*\*séwHnus or \*\*méntis, instead of \*suHnúš (p. 38) and \*mṇtís (p. 35) (with the accent marked as traditionally reconstructed) that are actually attested in the daughter languages but often considered as secondary without any real evidence.

The standard theory usually operates with the following basic and neatly organized ablaut/accent types: static (acrostatic/acrodynamic) \*nók<sup>w</sup>ts – gen. sg. \*nék<sup>w</sup>ts ‘night’ (immobile root accent); proterodynamic (proterokinetic) \*péh<sub>2</sub>wṛ – gen. sg. \*ph<sub>2</sub>wéns ‘fire’ (accent shifts from root to suffix); amphidynamic (amphikinetic, holokinetic/holodynamic) \*h<sub>2</sub>éwsōs – gen. sg. \*h<sub>2</sub>us(s)é/ós ‘dawn’ (accent shifts from root to ending); hysterdynamic (hysterokinetic) \*ph<sub>2</sub>tēr – gen. sg. \*ph<sub>2</sub>tré/ós ‘father’ (accent shifts from suffix to ending). However, here we shall use an analysis that is somewhat heterodox in comparison to the standard one because it stresses morpheme alternations and not the supposed PIE accent shifts, with emphasis on the last stage of PIE (and not some supposed pre-PIE “original” stage). In this analysis, the reconstructed ablaut types are somewhat different than those in the standard analysis, but the usual terminology is more or less preserved.

In our analysis here, we take it that all PIE morphemes (roots R, suffixes S, endings E) in a nominal paradigm were either strong (+), alternating (±), or weak (–). Strong morphemes always have the full (or lengthened) grade; alternating morphemes have the full (or lengthened) grade in some cases but the zero grade in others; weak morphemes are always in the zero grade. E.g., if we look at the roots of the already mentioned words, \*ne/ok<sup>w</sup>t- is (+), \*h<sub>2</sub>ews-/h<sub>2</sub>us- is (±), and \*suHn- is (–). Most endings were invariant (whether they were weak like nom. sg. \*-s or strong like nom. pl. \*-es) – only gen. sg. \*-es/-os/-s, instr. sg. \*-eh<sub>1</sub>/-h<sub>1</sub>, and perhaps nom. du. \*-h<sub>1</sub>/-h<sub>1</sub>e (cf. OLith. *-e* for the latter) had ablaut variants, i.e., were of the alternating type (thus, whether an ending in an ADT

is +, –, or ± is seen only in these two or three cases, while being irrelevant in others). The same morpheme did not have to be the same in all words – e.g., the suffix \*-i- was weak in \*potis and alternating in \*mētis (p. 35), etc. Thus, the (strong/alternating/weak) nature/valency of a morpheme was paradigmatic, not universal. A similar analysis can be applied to athematic verbs as well.

The following nominal ADTs can be reconstructed for the last stage of PIE:

- 1) **acrodynamic** (ἄκρος ‘outermost’, δύναμις ‘power’)
 

RSE + ± – \*me-h<sub>2</sub>tēr-Ø – gen. sg. \*me-h<sub>2</sub>tṛ-s ‘mother’
- 2) **proterodynamic** (πρότερος ‘in front’)
 

RSE ± ± – \*h<sub>3</sub>neh<sub>3</sub>-mṇ-Ø – gen. sg. \*h<sub>3</sub>ṇh<sub>3</sub>-men-s ‘name’
- 3) **mesodynamic** (μέσος ‘middle’)
 

RSE – ± – \*suHn-u-s – gen. sg. \*suHn-ow-s ‘son’
- 4) **amphidynamic** (ἀμφί ‘on both sides’)
 

RSE + – + \*pot-i-s – gen. sg. \*pot-y-e/os ‘master’ (cf. p. 77 for \*melit)
- 5) **hysterodynamic** (ὑστερος ‘coming after’)
 

RSE – ± + \*p-h<sub>2</sub>tēr-Ø – gen. sg. \*p-h<sub>2</sub>tr-e/os ‘father’
- 6) **holodynamic** (ὅλος ‘whole’)
 

type a) RSE + ± + \*h<sub>2</sub>ek-mō(n)-Ø – gen. sg. \*h<sub>2</sub>ek-mn-e/os ‘stone’  
 type b) RSE ± ± + \*pont-eH-s – gen. sg. \*pṇt-H-e/os ‘path’
- 7) **holostatic** (only neuter *s*-stems) (στατός ‘placed’)
 

RSE +++ \*neb<sup>h</sup>-os-Ø – gen. sg. \*neb<sup>h</sup>-es-e/os ‘sky’

What is relevant in an ADT are not only the types of morphemes but the way they pattern – e.g., two holodynamic types are put together because they differ only in the number of morphemes that alternate (but both have + in the nom. sg. and – in the gen. sg. in alternating morphemes), while acrodynamic (+ ± –) and proterodynamic (+ ± –) type are distinguished because they differ not only in what morphemes alternate, but also in how they alternate (cf. the zero suffix in the gen. sg. \*me-h<sub>2</sub>tṛ-s, but a full suffix in the gen. sg. \*h<sub>3</sub>ṇh<sub>3</sub>-men-s). Additional (sub)types may also be potentially added (cf. p. 77), and different naming/grouping of the types is always possible. In separate stems (e.g., *i-*, *n-*, etc.), one can usually find one to three types. Because of numerous analogies and divergent later developments, reconstructions are sometimes speculative, and different authors have different opinions. Some nouns may have had coexisting different ablaut variants. For slightly different ADTs in root nouns, cf. p. 71.

#### *Athematic case endings*

All athematic stems had mostly the same endings (differing in a number of cases from the *o*-stems), the differences being in the presence of \*-i- in the loc. sg. and neuter \*-h<sub>2</sub> in the nom./acc./voc. pl. (or synchronic \*-Ø). In the gen. sg. acro-, protero-, and mesodynamic stems had a zero-grade \*-s, while other types had a full-grade \*-es or \*-os. The

original distribution of \*-es and \*-os is not clear (it might have originally depended on accent, ADT, stem type, dialect, etc.) since most languages have generalized one of the variants – some languages have only \*-es (OCS *-e*), some have only \*-os (Gr. *-ος*), in some we cannot tell the original vowel (OInd. *-as*), and only rarely does one see both variants in one language (Lat. *-is* < OLat. *-es*, together with OLat. *-us* < \*-os). The nom. pl. \*-es is identical in form to the gen. sg. variant \*-es but is invariant (as are most case endings), unlike the gen. sg. ending (however, cf. the pronominal plural \*-s(-) – p. 83). In the gen. pl., one can assume an original short \*-om on structural grounds (with possible reflexes in Celtic, Italic, Anatolian, and Slavic), while most languages show reflexes of a long \*-ōm (presumably by early, perhaps partially already PIE, analogy to *o-* and *eh<sub>2</sub>-stems*).<sup>5</sup>

### *Compensatory lengthening in athematic stems*

A number of compensatory lengthening (CL) processes occurred word-finally in athematic nouns. The masculine/feminine ending \*-s dropped in pre-PIE after a root- or suffix-final resonant (\*m/n/l/r), \*y, and \*s, with subsequent CL (**Szemerényi's Law**): \*d<sup>h</sup>eg<sup>h</sup>ō(m) < \*-om-s (p. 77), \*h<sub>2</sub>ekmō(n) < \*-on-s (p. 76), \*h<sub>2</sub>ebōl < \*-ol-s (p. 77), \*ph<sub>2</sub>tēr < \*-er-s (p. 75), \*sek<sup>w</sup>Hōy < \*-oy-s (p. 73), \*h<sub>2</sub>ewsōs < \*-os-s (p. 79). The same kind of lengthening might have occurred in some other cases, like in \*kērd > \*kēr (with possible subsequent restoration of \*-d, p. 53). This lengthening was morphological, not a live phonological process, in the last stage of PIE as seen from forms like the gen. sg. \*dems (p. 71) and the acc. pl. ending \*-ons (p. 64). This law explains many *s*-less long-grade forms in the nom. sg. of athematic stems. However, the length in forms with the preserved \*-s, like \*pōds (p. 15), \*g<sup>w</sup>ōws (p. 24), \*(h<sub>2</sub>)nepōts (p. 39), \*-ōws (p. 74), and ptc. \*-ōnts (p. 104), is not completely clear. The original dropping of \*-s, CL, and then analogical restoration of \*-s is possible but perhaps not convincing (why would it be restored only in these forms and not after resonants, \*y, and \*s as well?). Perhaps it is a consequence of morphologization and analogical interplay of several earlier phonological processes like Szemerényi's Law, early monosyllabic lengthening (p. 54–55) in cases like \*pōds, and perhaps Stang's Law (see below) in the acc. sg. like \*g<sup>w</sup>ōm. In some cases, the length can be interpreted in more than one way, e.g. \*wīs 'poison' (p. 55) can be a result of either monosyllabic lengthening or CL from \*wis-s.

CL occurs with \*-h<sub>2</sub> after \*n, \*r, and \*s as well, cf. nom. pl. \*h<sub>3</sub>nh<sub>3</sub>mōn < \*-on-h<sub>2</sub> (p. 77), \*k<sup>w</sup>etwōr < \*-or-h<sub>2</sub> (p. 89), and nom. pl. \*neb<sup>h</sup>ōs < \*-os-h<sub>2</sub> (p. 79); and with \*-i after glides, cf. loc. sg. \*m̥tēy < \*-ey-i (p. 72) and loc. sg. \*suHnōw < \*-ow-i (p. 73). CL also occurred when glides and laryngeals dropped before the final \*-m (**Stang's Law**), cf. \*dyēm < \*-ewm (p. 40, 72), \*g<sup>w</sup>ōm < \*-owm (p. 71), and \*-ām < \*-eh<sub>2</sub>-m in thematic stems (p. 66).

### *Root nouns*

Most athematic nouns have a suffix of some kind between the root and endings and are usually disyllabic in the nom. sg. Those that do not have a suffix are monosyllabic in the nom. sg. and are called root nouns (the ending is added directly to the root), cf. \*pōd-s 'foot', gen. sg. \*ped-e/os. The root in root nouns can end in a stop (\*pōd-s), resonant (\*dō(m)-Ø 'home'), diphthong (\*dyēw-s 'sky'), cluster (\*kērd-Ø 'heart'), etc. According to the final segment of the root, root nouns can also be considered plosive stems (\*pōd-s), *m*-stems (\*dō(m)-Ø), etc. Root nouns had a tendency to become less frequent

or disappear altogether in many IE languages, being replaced with various suffixal derivatives. In some cases, it is not clear whether certain nouns are indeed root or suffixal stems; e.g., is \*kwō(n) ‘dog’ to be analyzed as a root noun or as an *n*-stem \*kw-ō(n) with a vowelless root (p. 41)? Root nouns had the following ablaut types (although they have no suffixes, root nouns can tentatively be fitted into the above ADT scheme):

- a) RE +- (acrodynamic) \*dō(m)-Ø – gen. sg. \*dem-s ‘home’
- b) RE ++ (amphidynamic) \*pōd-s – gen. sg. \*ped-e/os ‘foot’
- c) RE -+ (hysterodynamic) \*muHs-(s) – gen. sg. \*muHs-e/os ‘mouse’
- d) RE ++ (holodynamic) \*dyēw-s – gen. sg. \*diw-e/os ‘day sky’

Acrodynamic nouns had \*o in the nom./acc. and \*e in other cases, cf. \*dō(m) (Gr. Hom. δῶ, Arm. *tun*) – gen. sg. \*dems (Ved. *dán*, GAv. *dāng*, Gr. δεσ-πότης ‘master’ < \*dems potis ‘lord of the house’), or \*g<sup>w</sup>ōws ‘cow’ (Ved. *gáus*, Gr. βοῦς) – gen. sg. \*g<sup>w</sup>ews (Ved. *gós*, Av. *gāuš*). For the length in \*dō(m) or \*kwō(n) cf. p. 70. Stang’s Law operates in the acc. sg. \*g<sup>w</sup>owm > \*g<sup>w</sup>ōm (Ved. *gām*, Dor. βῶν, Umbr. *bum*) – cf. p. 70. The amphidynamic stems had a more complex ablaut pattern:

TABLE 1.17 IE ROOT NOUNS (SINGULAR)

sg.	PIE	Ved.	Gr.	Lat.
nom.	*pōd-s	<i>pát</i>	π\ού\ς	<i>p\ē\ś</i>
acc.	*pod-ṛ	<i>pāda m </i>	πόδα	<i>p\ē\dem</i>
gen.	*ped-e/os	<i>padás</i>	π\ο\δός	<i>pedis</i>
dat.	*ped-ey	<i>padé</i>	Myc. <i>po-de</i>	<i>pedī</i>
loc.	*ped-i	<i>padí</i>	[dat. π\ο\δί]	[abl. <i>pede</i> ]
instr.	*ped-eh <sub>1</sub>	<i>padá</i>		

Cf. p. 70 for the nom. sg. length.

TABLE 1.18 IE ROOT NOUNS (PLURAL)

pl.	PIE	Ved.	Gr.	Lat.
nom.	*pod-es	<i>pádas</i>	πόδες	<i>p\ē\dl\ē\ś</i>
acc.	*pod-ṛs	<i>padás</i> < *pádas <sup>a</sup>	πόδας	<i>p\ē\dl\ē\ś</i>
gen.	*ped-om	<i>pad\á\m</i>	π\ο\δ\ώ\ν	<i>pedum</i>
dat.	*ped-b <sup>h</sup> os	*padbh <sub>1</sub> yás		<i>ped ī bus</i>
loc.	*ped-su	<i>patsú</i>	[dat. π\ο\σσ\ί\]	
instr.	*ped-b <sup>h</sup> i(s)	<i>padbhís</i>	Myc. <i>p\o\pi</i>	

<sup>a</sup> Cf. RV *ápas* ‘waters’ (6x), together with the secondary *apás*.

Dual: nom./acc./voc. \*pod-h<sub>1</sub>(e) (Gr. πόδε), dat./instr./abl. \*ped-b<sup>h</sup>yoh<sub>1</sub> (Ved. *pad-bhyám*). The nom./acc. \*pod- (accented in Vedic/Greek) and oblique \*ped- stem are indicated by Ved. *pād-* (by Brugmann’s Law – p. 40, 206) and *pad-* respectively (Greek has generalized \*pod- and Latin \*ped-). The usual supposition is that the oblique \*ped- is secondary for the supposedly original holodynamic \*pd- (cf. Gr. ἔπι-βδ-α ‘the day after the festival’ < \*‘following the trace’, Ved. *upa-bd-ás* ‘trampling, rattle’, Av. *fra-bd-a-* ‘front foot’).

For holodynamic stems cf. \*dyēws (Ved. *dyáus*, Gr. Ζεύς ‘Zeus’), gen. sg. \*diwe/os (Ved. *divás*, Gr. Διός), acc. sg. \*dyēm with Stang’s Law (Ved. *dyám*, Gr. Ζήν, Lat. *diem* ‘day’), voc. sg. \*dyew! (Gr. Ζεῦ) or \*kwō(n) (Ved. *śwá*, Gr. κύων, Lith. *šuō*) – gen. sg. \*kune/os (p. 41, 56). Holodynamic stems show an \*e or \*o full grade in the nom./acc. and the zero grade in the oblique cases.

### *I*-stems

In the *i*-stems (as in the *u*-stems), the root was unchangeable (either full as in \*h<sub>3</sub>ewis ‘sheep’ or zero as in \*m̥tis ‘mind’), while the suffix could be in various ablaut grades (\*-i-/-ey-/-oy-/-ēy-). The *i*-stems had two ablaut types, mesodynamic and amphidynamic, of which the former generally prevailed in later IE languages. The mesodynamic *i*-stem (m./f.) paradigm in IE:

TABLE 1.19 IE *I*-STEMS (SINGULAR)

sg.	PIE	Ved.	Gr.	Lat.	OCS	Lith.	Goth.
nom.	*-i-s	- <i>is</i>	-ις	- <i>is</i>	- <i>b</i>	- <i>is</i>	- <i>s</i>
acc.	*-i-m	- <i>im</i>	-iv	- <i>em</i>	- <i>b</i>	- <i>i</i>	-∅
voc.	*-ey	- <i>e</i>	{-ι}	[- <i>is</i> ]	- <i>i</i>	- <i>iē</i>	
gen.	*-ey-s	- <i>és</i>	Hom. {-ηος}	{- <i>is</i> } <sup>a</sup>	- <i>i</i>	- <i>iēs</i>	- <i>a</i> vis <sup>b</sup>
dat.	*-ey-ey	- <i>aye</i>		- <i>i</i>	- <i>i</i>	{- <i>iai</i> }	{- <i>ai</i> }
loc.	*-ēy	- <i>á</i>	dial. -ηῖ		- <i>i</i>	{- <i>yjē</i> }	
instr.	*-i-h <sub>1</sub>	- <i>ī</i>			{- <i>bjō</i> }	{- <i>imi</i> }	

<sup>a</sup> Cf. Osc. original (and generalized) *-eis*; <sup>b</sup> cf. original OHG (f.) *-i* < PGmc \*-īz < PIE \*-eys.

For the invariant zero grade of the root cf. also \*ṅg<sup>(w)</sup>nis ‘fire’, \*k<sup>w</sup>rmis/\*w<sup>r</sup>mis ‘worm’ (p. 35), \*m̥tis ‘death’ (p. 35), and other nouns in \*-tis, nic. (presumably, only the *i*-stems with the zero-grade root originally had this type of endings). In the gen. sg., the zero-grade suffix was followed by a full-grade ending and vice versa (gen. sg. \*-y-e/os but \*-ey-s). Goth. gen. sg. (f.) *-ais* and ONor. *-ar* are probably secondary (by analogy to *u*-stem \*-auz < \*-ows), and there is no need to reconstruct a PIE gen. sg. variant \*-oys (OCS and OHG point to \*-eys). The loc. sg. \*-ēy is to be explained through a (pre-)PIE process of \*-ey-i > \*-ēy (p. 70).

TABLE 1.20 IE *I*-STEMS (PLURAL)

pl.	PIE	Ved.	Gr.	Lat.	OCS	Lith.	Goth.
nom.	*-ey-es <sup>a</sup>	- <i>ayas</i>	-εις	- <i>ēs</i>	- <i>bjē</i>	{- <i>ys</i> }	- <i>eis</i>
acc.	*-i-ns	- <i>īs</i> \	dial. -ivς	- <i>īs</i>	- <i>i</i>	- <i>is</i>	- <i>ins</i>
gen.	*-ey-om	{- <i>inām</i> }	-ε(ω)v	- <i>ium</i>	- <i>bi</i> <sup>c</sup>	- <i>iū</i> \	- <i>e</i>
dat.	*-i-b <sup>h</sup> /mos	- <i>ibh</i> [y]as		- <i>ibus</i>	- <i>bm̥b</i>	- <i>ims</i>	- <i>im</i>
loc.	*-i-su	- <i>īsu</i>	[dial. dat. -i σ ι\]		- <i>bx̥b</i>	- <i>y’is</i> { <i>ē</i> }	
instr.	*-i-b <sup>h</sup> i(s) <sup>b</sup>	- <i>ibhis</i>			- <i>bmi</i>	- <i>imis</i>	

<sup>a</sup> Cf. also the Hitt. generalized nom. pl. *-eš* < \*-eyes; <sup>b</sup> BSL. \*-miHs; <sup>c</sup> \*-bj̥b.

The full grade (\*-ey-) of the suffix is found in the gen./dat./loc./voc. sg. and nom./gen. pl. – other cases have the zero grade (\*-i-). Dual forms are nom./acc./voc. \*-i-h<sub>1</sub> (Ved. -ī, OCS -i), dat./instr./abl. \*-i-b<sup>h</sup>y/moh<sub>1</sub> (Ved. -ibhyām|, OCS -ьma), gen./loc. \*-ey-ow < \*-ey-u? (p. 63) (OCS *kost-bju* is likely archaic, since it is different from the other dual forms, while Ved. gen./loc. du. *hāryos* ‘tawny’ < \*-y-ow-s is probably analogical to them, though OCS form could also theoretically be due to analogy to the gen. pl.). The neuter forms differed in the nom./acc./voc. only: sg. \*-i (Ved. *vāri* ‘water’, Lat. *mare* ‘sea’, Hitt. *tuppi* ‘clay tablet’) and pl. \*-ih<sub>2</sub> (Ved. adj. *śúcī* ‘bright’, Hitt. *armizzi* ‘bridges’).

The scarcely attested amphidynamic paradigm had a weak suffix (\*-i-/-y-) in all cases, cf. gen. \*pot-y-e/os ‘master’ (Gr. πόσιος ‘husband’), \*h<sub>2</sub>ew-y-e/os ‘sheep’ (Ved. *ávyas*, Gr. Ion. διος), dat. \*-y-ey (Ved. *pátye* ‘husband’), loc. \*-y-i (Gr. dat. οἰί), voc. sg. \*-i (Gr. πόσι!), nom. \*-y-es (Ved. adj. *aryás* ‘loyal’, Gr. Hom. ὄϊες), gen. pl. \*-y-om (Gr. οἰών). One could assume that originally all *i*-stems with a full-grade root – like \*g<sup>h</sup>ostis ‘guest’ (p. 40), \*klownis ‘hip’ (p. 51), \*h<sub>3</sub>eg<sup>wh</sup>is ‘snake’ (Ved. *áhis*, Gr. ὄφις ‘serpent’), etc. – were amphidynamic (i.e., that the ablaut of the suffix was connected to the ablaut of the root), even when we have no direct attestation for that (unlike the case in \*h<sub>3</sub>ewis, \*potis), though acrodynamic + ± – stems like \*\*g<sup>h</sup>ost-ey-s (perhaps later and analogical) are not unimaginable. A holodynamic + ± + type can also be reconstructed, cf. \*-ōy < \*-oy-s (Ved. *sákhā* ‘friend’, Gr. ἠχώ ‘echo’, Hitt. *zahḫaiš* ‘battle’), acc. sg. \*-oy-ṛṇ (Ved. *sákhāyam*, Hitt. *zahḫaiš*), voc. sg. \*-oy (Gr. Σαφοῖ!), gen. sg. \*-y-e/os (Ved. *sákhya* with secondary -ur, Hitt. *zahḫiyaš*), etc. For \*-ih<sub>2</sub>(s) f. stems see p. 67.

*U-stems*

The *u*-stems were mostly parallel to the *i*-stems concerning ADTs. The root was invariant in masculine/feminine stems but not in all neuter stems (cf. \*dor-u – gen. sg. \*dr-ews below). The mesodynamic (m./f.) paradigm prevailed in later languages:

TABLE 1.21 IE U-STEMS (SINGULAR)

sg.	PIE	Ved.	Gr.	Lat.	OCS	Lith.	Goth.
nom.	*-u-s	-ús	-υς	-us	-ь	-ùs	-us
acc.	*-u-m	-úm	-υv	-um	-ь	-ų	-u
voc.	*-ow	-o		{-us}	-u	-aũ	-au
gen.	*-ow-s	-ós <sup>a</sup>	Hom. {-εος}	-ūs <sup>c</sup>	-u	-aũs	-aus
dat.	*-ew-ey	-áve	{-ει}	-uī	-ovi	{-uī}	-a <u>u</u> <sup>d</sup>
loc.	*-ōw	-au <sup>b</sup>			-u	{-ujė}	
instr.	*-u-h <sub>1</sub>	GA <i>v.</i> -ū			{*-ьmь}	{-umì}	

<sup>a</sup> GA*v.* -āuš; <sup>b</sup> *mán-au* ‘man’; <sup>c</sup> OLat. SENAT-OVS ‘of the Senate’, Osc. *castr-ous* (cf. Lat. *castrum* ‘fort’); <sup>d</sup> early runic/early OHG -iu < PGmc \*-iwi < PIE \*-ewey.

For the invariant zero grade of the root cf. also \*p<sup>h</sup>rtus (Av. *pərtu-* ‘crossing’, Lat. *portus* ‘harbor’, OEng. *ford*), \*t<sup>h</sup>h<sub>2</sub>us ‘thin’ (p. 16), \*g<sup>w</sup>ṛh<sub>2</sub>us ‘heavy’ (p. 39) and other adjectives, etc. Loc. \*-ōw derives from pre-PIE \*-ow-i (like *i*-stem loc. sg. \*-ēy above). Loc. \*-ēw is also reconstructable (cf. \*-ēy in *i*-stems), but the non-palatal -u (i.e., not -ju) in OCS *synu* is easier to explain from \*-ōw (cf. the \*-eys : \*-ows difference between *i*- and *u*-stems in the gen. sg. as well).

TABLE 1.22 IE *U*-STEMS (PLURAL)

pl.	PIE	Ved.	Gr.	Lat.	OCS	Lith.	Goth.
nom.	*-ew-es	-ávas	dial. -εεζ	{-ūs}	-ove	{-ūs} <sup>c</sup>	-jus
acc.	*-u-ns	-ávn	dial. -υυζ	-ūs	-y	-us	-uns
gen.	*-ew-om	{-ūnám}	-ε(ω)υ	-uum	-ovъ	{-ū}	-iw{e}
dat.	*-u-b <sup>h</sup> /mos	-úbh y as		-ibus	-ьmb	-ūms	-um
loc.	*-u-su	-úšu	<sup>b</sup>		-ьxb	dial. -us\è\	
instr.	*-u-b <sup>h</sup> i(s) <sup>a</sup>	-úbhis			-ьmi	-umis	

<sup>a</sup> BSl. \*-miHs; <sup>b</sup> dat. pl. δάκρυ(σ)|\ 'tears'; <sup>c</sup> dial. -aus < \*-awes < PIE \*-ewes.

The full grade (\*-ew-/ow-) of the suffix is found in the gen./dat./loc./voc. sg. and nom./gen. pl. (the \*e-/o-grade distribution is different from in *i*-stems) – other cases have the zero grade (\*-u-). Dual forms are nom./acc./voc. \*-u-h<sub>1</sub> (Ved. -ū, OCS -y, Lith. -u), dat./instr./abl. \*-u-b<sup>h</sup>y/moh<sub>1</sub> (Ved. -ubhyā|m|, OCS -ьma), gen./loc. \*-ew-ow < \*-ew-u? (p. 63) (cf. OCS *synovu* < \*-ew-ow but probably secondary Ved. *bāhvós* ‘arms’ < \*-w-ow-s, like in *i*-stems, p. 73). Neuters had a nom./acc./voc. sg. \*-u (Ved. *mádhū* ‘sweetness’, Gr. μέθυ ‘wine’, Goth. *faihu* ‘cattle’, Hitt. *gēnu* ‘knee’), nom./acc./voc. pl. \*-u-h<sub>2</sub> (Ved. *vásū* ‘wealths’, Hitt. *āššū* ‘goods’), and nom./acc./voc. du. \*-w-ih<sub>1</sub> (Ved. adj. *urvī* ‘wide’).

The vestiges of the amphidynamic type (with weak suffix \*-u-/w- in all cases) are often attested in neuters (e.g., \*med<sup>h</sup>u ‘honey’, \*peku ‘cattle’): gen. \*-w-e/os (Ved. *mádhvas*, *krátvas* (m.) ‘ability’, Gr. Myc. *me-tu-wo*, older Lat. *senatuos* ‘of the Senate’), dat. \*-w-ey ‘cattle’ (Ved. *pásve*, Goth. *mann* (m.) < \*manw- ‘man’), loc. sg. \*-w-i (Gr. Hom. dat. γουί < \*γouví ‘knee’), nom. \*-w-es (YAv. *pasuuō*), and gen. pl. \*-w-om (YAv. *pasuuqm*). It can be perhaps be assumed that all *u*-stems with full-grade roots were originally amphidynamic (even in cases where there is no direct evidence for that), e.g. \*génus ‘jaw’ (Ved. *hánus*, Gr. γένυς) – gen. sg. \*génwe/os (Goth. *kinnus* ‘cheek’ with generalized -nn- < \*-nw-), adj. \*h<sub>1</sub>oh<sub>1</sub>kus ‘fast’ (Ved. *āsús*, Gr. ὠκύς), etc. Cf. also the proterodynamic neuter \*dor-u-Ø ‘tree’ (Ved. *dāru*, Gr. δόρυ, Hitt. *tāru*) – gen. sg. \*dr-ew-s (Ved. *drós*, YAv. *draoš*), though this might not be the oldest/only ablaut type; cf. the ablaut variant \*derw- in OCS *drěvo* < \*dervo ‘tree’, Lith. *dervà* ‘resin, tar’. The original ablaut pattern of \*gén-u- (Lat. *genū*, Hitt. *gēnu*)/\*gón-u- (Ved. *jānu*, Gr. γόνυ)/\*gñ-ew- (Hitt. *ganu-*, Goth. thematized *kniu*) ‘knee’ is controversial (see above for Gr. dat. γουί). There is some evidence of a holodynamic type with a nom. sg. in \*-ē/ōws (m./f.): Av. *hiθāuš* ‘associate’, Gr. ἵππεύς ‘horseman’, πάτρως ‘paternal uncle’, Hitt. *harnauš* ‘birthing chair’.

The suffix \*-u- is found in a feminine motion ending \*-u-h<sub>2</sub>-s as well, cf. \*swekuros ‘father-in-law’ (Ved. *śvāsuras*, Gr. ἐκυρός, Lat. *socer*, Lith. *šėšuras*) and \*swekruh<sub>2</sub>s ‘mother-in-law’ (Ved. *śvasrūs*, OCS *svekry*), declined like a normal consonant amphidynamic stem (gen. sg. \*swekruh<sub>2</sub>e/os > OCS *svekrъve*, cf. Ved. *tanúas* ‘body’) and root nouns with roots ending in \*-H- like \*h<sub>3</sub>b<sup>h</sup>ruHs ‘eye-brow’ (p. 48) or \*suHs ‘swine’ (p. 38).

### *R*-stems

The *r*-stems are dominantly masculine/feminine nouns with an invariant root. Many of them belong to nouns denoting family relations – \*ph<sub>2</sub>tēr ‘father’ (p. 46), \*meh<sub>2</sub>tēr ‘mother’ (p. 45), \*d<sup>h</sup>ugh<sub>2</sub>tēr ‘daughter’ (p. 47), \*b<sup>h</sup>reh<sub>2</sub>tēr ‘brother’ (p. 18), \*swesōr ‘sister’ (p. 32), \*deh<sub>2</sub>iwēr ‘husband’s brother’ (Ved. *devár-*, Lith. *dieveris*, OCS *děverb*),

\*ǵenh<sub>1</sub>tōr ‘begetter’ (p. 46), etc. *Nomen agentis* \*-tēr/\*-tōr nouns were also *r*-stems (cf. PIE \*dh<sub>3</sub>tēr > Gr. δοτήρ but \*deh<sub>3</sub>tōr > Gr. δώτωρ, Ved. *dātā*; Lat. *dātor*, gen. sg. *-ōris* is a mix of the two types – all ‘giver’). Some monosyllabic words ending in \*-r# (like \*h<sub>2</sub>stēr ‘star’, \*h<sub>2</sub>nēr ‘man’) can be treated as root nouns – in such cases it is sometimes impossible to tell if a segment was originally part of the suffix or the root. One can reconstruct two basic ablaut types for most *r*-stems – hystero-dynamic (H) and acro-dynamic (A) – that differed only in a few cases:

TABLE 1.23 IE R-STEMS (SINGULAR)

sg.	PIE	Ved.	Gr.	Lat.	OCS	Lith.	Goth.
nom.	*-ēr	-á	-ηρ/-ήρ	-er	-i	-ėl-ė̃	-ar
acc.	*-er-ṛ	-ára m	-έρα	-v <sup>h</sup> em	-erь	-erį	-ar
voc.	*-er	-ar	-ερ	-er	[-i]	{-erį̃}	[-ar]
(A) gen.	*-r-s	-úr <sup>a</sup>					<sup>d</sup>
(H) gen.	*-r-e/os	<sup>b</sup>	-ρός	-ris	- e re	- e r̃s <sup>c</sup>	-rs
dat.	*-r-ey	-ré		-rī	- e rī	{-erįai}	-r
(A) loc.	*-r-i		[dat. -pī]		{-ere}		
(H) loc.	*-er-i	-ári				{-erįjė̃}	
(A) instr.	(?) *-r-h <sub>1</sub>				{-erьjǫ}	{-erimi}	
(H) instr.	*-r-eh <sub>1</sub>	-rā					

<sup>a</sup> Av. *n-ər<sup>3</sup>š* ‘man’; <sup>b</sup> Av. *brāθ-rō* < \*-rah ‘brother’; <sup>c</sup> OLith./dial. *-erės*; <sup>d</sup> ONor. *fǫd-ur*, OEng. *fad-ur*, *fead-ur* < PGmc \*-urz.

The length in the nom. sg. is due to Szemerényi’s Law (p. 70). The vowel \*e is more frequent in the nom. sg. than \*o.

TABLE 1.24 IE R-STEMS (PLURAL)

pl.	PIE	Ved.	Gr.	Lat.	OCS	Lith.
nom.	*-er-es	-áras	-έρες	{-rēs}	-ere	-ers <sup>c</sup>
(A) acc.	*-r-ns	-r̃ñ <sup>a</sup>				
(H) acc.	*-er-ṛs		-έρας	-v <sup>h</sup> ēs	-eri	-eris
gen.	*-r-om	{-r̃nām <sup>b</sup> }	-ρῶν	-rum	- e r̃	{-erį̃}
dat.	*-r-b <sup>h</sup> /mos	-r̃bh y as		-r ī bus	{-erьmь}	{-erims}
loc.	*-r-su	-r̃su	[dat. -ράσι]		{-erьxь}	{-erįsė̃}
instr.	*-r-b <sup>h</sup> i(s)	-r̃bhis			{-erьmi}	{-erimis}

<sup>a</sup> GAv. *māt-ərəš* < \*-r-ns (p. 31); <sup>b</sup> cf. Ved. *svás-rām* ‘sisters’, *ná-rām* ‘men’; <sup>c</sup> OLith./dial. *-eres*.

The PIE ablaut pattern of the suffix is reconstructed mostly from the overlap of Vedic and Greek patterns, which look original (the ablaut is mostly generalized in other languages). The acro-/hystero-dynamic opposition is seen in different attested endings in the daughter languages in the gen. sg., in the difference between Vedic and Greek in the loc. sg., in structural analogy to the gen. sg. in the loc. sg., and in the difference between Vedic and Greek in the acc. pl. From the reflexes, it is impossible to deduce which nouns had which type. For structural reasons, one can assume that the derivatives with an invariant full-grade root (\*meh<sub>2</sub>tēr, \*b<sup>h</sup>reh<sub>2</sub>tēr, \*swesōr, \*deh<sub>2</sub>iwēr, \*ǵenh<sub>1</sub>tōr,



etc.) were acrodynamic (full-grade root + zero-grade ending in gen./instr. sg.), while the nouns with an invariant zero-grade root (\*ph<sub>2</sub>tēr, \*d<sup>h</sup>ugh<sub>2</sub>tēr, etc.) were hysterodynamic (zero-grade root + full-grade ending). An exceptional proterodynamic type with ablauting root can tentatively be reconstructed for ‘husband’s brother’s wife’: \*yen-h<sub>2</sub>tēr (Gr. ἐνάρηρ, OLith. *jentē*) – gen. sg. \*yη-h<sub>2</sub>ter-s (Skr. *yātār-*).

### *N-stems*

Masculine/feminine *n*-stems point to a holodynamic + ± + type with a complex suffixal ablaut pattern:

**TABLE 1.25 IE *N*-STEMS (SINGULAR)**

sg.	PIE	Ved.	Gr.	Lat.	OCS	Lith.	Goth.	Hitt.
nom.	*-ō(n)	-ā	-ων	-ō	- y	-uō	- a  <sup>f</sup>	-a š
acc.	*-en-η <sup>a</sup>	-āna m	- o va	-inē	-enь	-enī	- a n	- a nan
voc.	*-on	-an	-ov <sup>c</sup>	[-ō]	[-y]	{-eniē}		
gen.	*-n-e/os	-nas	- o voς	- i nis	- e ne	- e nīs <sup>e</sup>	- i ns	- a naš
dat.	*-n-ey	-ne <sup>b</sup>		- i nī	- e ni	{-eniui}	{-in}	
loc.	*-en	-an	<sup>d</sup>		-en e	{-enyjē}		<sup>g</sup>
instr.	*-n-eh <sub>1</sub>	-nā			{-enьmь}	{-enimi}		

<sup>a</sup> With a younger variant \*-on-η; <sup>b</sup> *mahim-né* ‘greatness’; <sup>c</sup> δαῖμ-ov! ‘god’; <sup>d</sup> adv. αἰών ‘always’ < \*h<sub>2</sub>eywen (cf. Gr. αἰών ‘age’ < \*h<sub>2</sub>eywō(n)); <sup>e</sup> older *-enes*; <sup>f</sup> cf. *-o* (f./n.); <sup>g</sup> loc. sg. (n.) *laman* ‘name’.

The length in \*-ō(n) < \*-on-s is due to Szemerényi’s Law (p. 70). Some believe that the final \*-n had dropped already in PIE (i.e., \*h<sub>2</sub>ekmō, \*kwō ‘dog’ unlike \*swesōr ‘sister’), cf. Lat. *-or*, OIr. *-ur* < \*-ōr in *r*-stems but Lat. *-ō*, OIr. *-u* < \*-ō(n) in *n*-stems (unlike Gr. *-ων*, Arm. *-un* < \*-ōn, and OCS *-y* < \*-ōn-|s|, which would then have reintroduced \*-n-(s) by analogy later). Perhaps it is best to assume that both variants (\*-ōn and \*-ō) occurred in PIE, just as in \*-ō(m) in *m*-stems (p. 77). Besides the usual \*-ō(n), there might have also existed a less frequent \*-ē(n), cf. the Greek ποιμήν ‘herdsman’ (but Lith. *piemuō*) – see below. As in *s*-stems (p. 78), the loc. sg. points to a zero-ending (\*-en-Ø). This is different from the *i*- and *u*-stems, where \*-ēy and \*-ōw can be derived from older \*-eyi and \*-owi and thus do not have an original zero-ending.

**TABLE 1.26 IE *N*-STEMS (PLURAL)**

pl.	PIE	Ved.	Gr.	Lat.	OCS	Lith.	Goth.	Hitt.
nom.	*-on-es	-ānas	-ovες	-in ē s	- e ne	- e ns	-ans	-aniš
acc.	*-en-ηs	-anas	- o vaς	-inēs	-eni	-enis	{-ans}	- a n u š <sup>b</sup>
gen.	*-n-om	-n ā m <sup>a</sup>	- ō v o v	- i num	- e nь	{-eny}	- a n e	
dat.	*-η-b <sup>h</sup> /mos	-abh y as		{-inibus}	{-enьmь}	{-enims}	{-am}	
loc.	*-η-su	-asu	[dat. - o σ t ]		{-enьxь}	{-enyšē}		
instr.	*-η-b <sup>h</sup> i(s)	-abhis			{-enьmī}	{-enimī}		

<sup>a</sup> *dur-nāmnām* ‘ill-named’; <sup>b</sup> *argamanuš* ‘tributes’.

The original ablaut of the suffix (the interchanging \*-on-/-en-/-n-) is best preserved in Vedic, with the other languages usually generalizing one ablaut grade in most forms (e.g., \*-on- in Greek). Generalized BSl. oblique \*-en- (which can hardly stem from the loc. sg. \*-en alone) and Ved. nom. pl. *-ānas* but acc. pl. *-anas* could point to the same original ablaut variation in the sg. as well (cf. Lith. *-uō* but *-eni*). It seems that some nouns had an \*-en- in all (full-grade) cases, e.g., (hysterodynamic) \*uk<sup>w</sup>s-ēn ‘ox’ – nom. pl. \*uk<sup>w</sup>s-en-es (Ved. *ukṣānas*, OEng. *exen* ‘oxen’). Some possible root nouns (like \*kwō(n) ‘dog’) can be treated as *n*-stems as well (‘dog’ can originally be \*p<sup>k</sup>w-ōn from \*peku ‘cattle’ – p. 52). In the neuter *n*-stems, the word ‘name’ and other derivatives in \*-men- point to a proterodynamic declension: \*h<sub>3</sub>neh<sub>3</sub>-m<sub>ṇ</sub> (Ved. *nāma*, Lat. *nōmen*, Gr. ὄνομα) – gen. sg. \*h<sub>3</sub>ṇh<sub>3</sub>-men-s (OIr. gen. sg. *anmae*; GA. *haxmāṅg* < \*-anh < \*-ens, Goth. *-ins* < \*-en-es < \*-en-s; OCS *im-*, OPruss. *emm-* < BSl. \*inHm-) – nom. pl. \*h<sub>3</sub>n(e)h<sub>3</sub>-mōn (Ved. *nāmā*). In the neuter nom. pl., one has to suppose pre-PIE \*-on-h<sub>2</sub> > \*-ōn with a compensatory lengthening (p. 70). For the \*-m<sub>ṇ</sub> (gen. sg. \*-men-s) derivatives cf. also Ved. *kār-ma* ‘act’ (⇒ Eng. *karma*), Gr. θεμα (⇒ Eng. *theme*), Lat. *sē-men* ‘seed’ (⇒ Eng. *semen*), etc.

### *M-, l-, laryngeal, and plosive stems*

Unlike the frequent *r*- and *n*-stems, other resonant stems were less frequent. For *m*-stems, cf. the root noun \*dō(m) (p. 71) and the holodynamic (type ±±±) \*d<sup>h</sup>eḡ<sup>h</sup>-ō(m) ‘earth’ (Hitt. *tēkan*, cf. p. 53) – gen. sg. \*d<sup>h</sup>ḡ<sup>h</sup>-m-e/os (Hitt. *taknāš* with an analogical *-n-*; Gr. χθών has the gen. root and the nom. suffix) and \*ḡ<sup>h</sup>ey-ō(m) ‘winter’ (Gr. χιών ‘snow’, Arm. *jiwn* ‘snow’, YAv. *ziia* all with secondary \*ḡ<sup>h</sup>i-, cf. OCS *zim-a*, Gr. χειμών for \*ḡ<sup>h</sup>ey-m-) – acc. sg. \*ḡ<sup>h</sup>y-em-ṇ (Lat. *hiem-*) – gen. sg. \*ḡ<sup>h</sup>i-m-e/os (GA. gen. sg. *zimō*) (p. 22). Like in *n*-stems (p. 76), the final \*-m was unstable after a long vowel in the nom. sg. (cf. Hom. δῶ ‘house’ < PIE \*dō(m) but otherwise -ow; Lat. *homō* ‘man’ but Hitt. *tēkan*; Arm. *jiwn*, etc.).<sup>6</sup> In *l*-stems, one finds just a (potentially root noun) \*seh<sub>2</sub>l-s ‘salt’ with a controversial vocalism/ablaut pattern (p. 42) and (dialectal?) \*h<sub>2</sub>ebōl ‘apple’ (Lith. *óbuol-as*) (p. 16).

Of the rare laryngeal stems, the holodynamic (±±±) word for ‘path’ is easily reconstructable (with Vedic preserving the original ablaut pattern): \*pont-e/oH-s (Ved. *pánthās*, cf. generalized \*pont- in Gr. πόντος ‘sea’, Lat. *pont-* ‘bridge’, OCS *poṭb-*) – gen. sg. \*p<sup>ṛ</sup>ṭ-H-e/os (Ved. *pathās*, cf. generalized \*p<sup>ṛ</sup>ṭ- in Gr. πῆτος, OPruss. *pintis*). An early proterodynamic laryngeal stem, transformed in most descendant languages to an *eh<sub>2</sub>*-stem (p. 24, 66), was \*g<sup>w</sup>en-h<sub>2</sub> ‘woman’ (OIr. *bé*, Ved. *jāni-*) – gen. sg. \*g<sup>w</sup>n-eh<sub>2</sub>-s (OIr. gen. sg. *mná*, Ved. nom. sg. *gnās* ‘divine woman’, cf. *gnās-pāti-* ‘husband of a divine woman’ for the gen. sg.). Cf. also the adjective \*meḡ-h<sub>2</sub>-s (p. 53, 79–80).

Stems ending in a stop were rare (cf. the root noun \*pōd-s – p. 15, 60, 70–71). The masculine word \*(h<sub>2</sub>)nepōts ‘grandson, nephew’ (Ved. *nápāt*, Lat. *nepōs* – p. 39) is widely attested, though it is unclear whether it was acrodynamic (gen. sg. \*(h<sub>2</sub>)nept-s > Ved. *náptur* with secondary \*-ṛs from *r*-stems) or ±±± holodynamic (gen. sg. \*(h<sub>2</sub>)nep-t-e/os > YAv. *naptō*). The neuter word \*mel-it ‘honey’ (Hitt. *milit*, Gr. μέλι, Goth. acc. sg. *milib*) with the \*-it- suffix (cf. Hitt. *šepitt-* ‘a kind of grain’) perhaps had a special ±-± amphidynamic subtype (Hitt. dat./loc. sg. *malitti* < \*ml-it-ey, cf. CLuw. generalized *mallit-*).

### *Heteroclitc stems*

Heteroclitcs were an ancient class of PIE neuter nouns with suppletive suffixes (one variant in the nom./acc., the other in the oblique cases). Though sometimes difficult to reconstruct exactly (with many languages often showing only isolated relics), their number is

not so small in total. The most common suppletion type was \*-r-/-n-: cf. Ved. *ás-ṛ-k* – gen. sg. *as-n-ás*, Hitt. *ēšḫ-ar* – gen. sg. *išḫa-n-āš* ‘blood’, Hitt. *wāt-ar* – gen. sg. *wit-en-aš* ‘water’, etc. The suppletion \*-l-/-n- is attested in the word ‘sun’ (GAv. *huuarē* – gen. sg. *x<sup>n</sup>āng* < \*sh<sub>2</sub>wens), and there may have been some other, less frequent and more controversial, types of suppletion (cf. p. 429, 437). Some languages in some cases preserve the actual suppletion as in PIE: Hittite abundantly (cf. Hitt. *paḫḫu-r* – gen. sg. *paḫḫu-en-aš* ‘fire’, p. 15), often Indo-Iranian (cf. Ved. *yák-ṛ-t* – gen. sg. *yak-n-ás* ‘liver’), sometimes Greek (ἥπ-αρ – gen. sg. ἥπ-α-τος ‘liver’ with -α- < \*-η-), and Latin (*iec-ur* – gen. sg. *ioc-in-eris* ‘liver’, p. 33). Most descendant languages usually generalize one of the variants (e.g., ONor. *vatn* but OEng. *wæter* ‘water’; Lat. *sōl*, Goth. *sauil* but Goth. also *sunno* ‘sun’; Goth. *fon* but OEng. *fȳr* ‘fire’; CS *ikra* ‘roe’ but OLith. *jėknos* ‘liver’). PIE nom. sg. forms are usually reconstructable (\*h<sub>1</sub>esh<sub>2</sub>r̥ ‘blood’, \*wod<sub>ṛ</sub> and coll. \*wedōr ‘water’, \*peh<sub>2</sub>w<sub>ṛ</sub> ‘fire’, \*yēk<sup>w</sup><sub>ṛ</sub> ‘liver’), though not always (cf. \*seh<sub>2</sub>-w-l- ‘sun’ with a problematic nom. sg. ablaut pattern, often reconstructed as \*seh<sub>2</sub>wōl – p. 32), while the reconstruction of oblique cases and ADTs is often more speculative.

### S-stems

Numerous neuter *s*-stems (with the \*-e/os- suffix) had a unique holostatic ADT with all full grades (+++), like the thematic stems (p. 62) and unlike other athematic stems. Neuter *s*-stems are considered athematic only because they have no thematic vowel, but their characteristics, such as the invariant root (though this occurs in other athematic stems as well) and only qualitative (\*-e-/o-) ablaut alternation of the suffix, remind one more of the thematic *o*-stems than of the other athematic stems:

TABLE 1.27 IE S-STEMS (SINGULAR)

sg.	PIE	Ved.	Gr. (Hom.)	Lat.	OCS	Hitt.
nom./acc./voc.	*-os	-as	-ος	-us	-\o\	-\āš
gen.	*-es-e/os	-asas	-εος b	-eris	-ese	-išaš
dat.	*-es-ey	-ase		-erī	-esi	
loc.	*-es <sup>a</sup>	-as i	[dat. -ε i ]	[abl. -er e ]	-es e	-iš
instr.	*-es-eh <sub>1</sub>	-asā			{-esъmъ}	

<sup>a</sup> Younger analogical form: \*-es-i; <sup>b</sup> Myc. *we-te-e* ‘year’.

The original loc. sg. has \*-Ø (cf. p. 76 for *n*-stems).

TABLE 1.28 IE S-STEMS (PLURAL)

pl.	PIE	Ved.	Gr. (Hom.)	Lat.	OCS
nom./acc./voc.	*-ōs	-ā m s i		{-era}	{-esa}
gen.	*-es-om <sup>a</sup>	-as\āvm		{-erum}	-esъ
dat.	*-es-b <sup>b</sup> /mos	-obh y as <sup>b</sup>		-er i bus	{-esъmъ}
loc.	*-e(s)-su	-a(s)su	[dat. -εσ(σ)\i\]		{-esъxъ}
instr.	*-es-b <sup>hi</sup> (s)	-obhis <sup>c</sup>			{-esy}

<sup>a</sup> Cf. Hitt. *nepišan*; <sup>b</sup> < \*-asbhyas; <sup>c</sup> < \*-asbhis; <sup>d</sup> Myc. *pa-we-pi* ‘pieces of cloth’.

The original nom. pl. \*-ōs (from older \*-os-h<sub>2</sub>, p. 70), which cannot be explained as secondary, can be seen regularly in Avestan (cf. *raocā* ‘lights’ < \*-āh), contaminated in Ved. *-āmsi*, and indirectly perhaps in OEng. *lombor* ‘lambs’ (< PGmc \*-ōz-a). Reflexes of the nom. pl. \*-es-eh<sub>2</sub> etc. are secondary. The distribution of suffixal \*-os-/-es- reminds of the \*pod-/ped- distribution in the root nouns (p. 71).

Masculine/feminine (holostatic) derivatives in \*-ēs (derived from neuter *s*-stems) also existed – cf. PIE \*h<sub>1</sub>su-men-ēs ‘well-disposed’ (Ved. *sumānās*, Gr. εὐμενής) from \*menos – gen. sg. \*menese/os (Ved. *mānas* ‘mind’, Gr. μένος ‘spirit’), Greek names ending in -κλήης (like Sophocles), or Lat. *Cerēs* (goddess). A holodynamic (±±±) feminine noun can be reconstructed: \*h<sub>2</sub>ews-ōs-Ø (< \*-os-s, p. 70) ‘dawn’ (Gr. Aeol. αὔως, Lat. generalized *aurōr-*; Ved. *uśās* and GAv. *uśā* with analogical *u-*) – gen. sg. \*h<sub>2</sub>us-s-e/os (Ved. *uśās*); cf. also Lat. *honōs* ‘honor’ – gen. sg. *honōris* with generalized -ōs- for the type. Note also the perf. ptcp. in \*-wos- (p. 104). Other, shorter and infrequent, *s*-stem types can perhaps be reconstructed as well, cf. Ved. *ākṣi* < \*h<sub>3</sub>ek<sup>w</sup>-s-Ø (neuter) with OCS *oko* – gen. sg. *očese* ‘eye’ (like *nebo*). There are also root nouns ending in (originally suffixal?) \*-s, e.g., \*muHs (p. 38, 71).

## ADJECTIVES

PIE adjectives were morphologically identical to nouns. However, the number of possible stems in adjectives was much smaller. For the most numerous *o*-stems cf. \*new-o-s (m.) – \*new-e-h<sub>2</sub> (f.) – \*new-o-m (n.) ‘new’ (Ved. *nāvas* – *nāvā* – *nāvam*, Gr. νέος – *vēā* – *véon*, Lat. *nouus* – *noua* – *nouum*, OCS *novъ* – *nova* – *novo*, Lith. *naĩ|j|as* – *naulj|à*, p. 32). For the frequent *u*-stems cf. \*sweh<sub>2</sub>dus (m.) – \*sweh<sub>2</sub>dwh<sub>2</sub> (f.) – \*sweh<sub>2</sub>du (n.) ‘sweet’ (Ved. *svādús* – *svādvī* – *svādú*, Gr. ἡδύς – ἡδ|ε|ῖα – ἡδύ; cf. Lith. *saldūs* – *salđì* ‘sweet’; Goth. *kaiúrus* ‘heavy’). Note that the \*-us form is used for the masculine only, unlike in nouns (p. 73). For the apparently not-so-frequent *i*-stems cf. \*semh<sub>2</sub>l-is (m./f.) – \*semh<sub>2</sub>l-i (n.) ‘similar’ > Lat. *similis* (m./f.) – *simile* (n.), OIr. *samail* ‘likeness’ (f.) (cf. also Hitt. *palh-i-š* ‘wide’). Original root adjectives may perhaps be reconstructed when dealing with different suffixes in later languages, i.e., \*nog<sup>w</sup>-s ‘naked’ (cf. Ved. *nagnás*, Gr. γυμνός, Lith. *niogas*, OCS *nagъ*, Goth. *naqaps*, Hitt. *nekumanza*). A ±±± holodynamic laryngeal stem can be reconstructed for \*meg<sup>w</sup>-h<sub>2</sub>-s (m./f.) – \*meg<sup>w</sup>-h<sub>2</sub> (n.) ‘big’ (Gr. μέγας (m.) – μέγα (n.), Ved. *māhi* (n.), Hitt. *mēk* (n.)), gen. sg. \*mǵ<sup>w</sup>-h<sub>2</sub>-e/os (Gr. Hom. ἄγα- ‘greatly-’, acc. sg. \*mǵ<sup>w</sup>-eh<sub>2</sub>-m (Aeol./Dor. adv. ἄγαν ‘very much’). Cf. p. 79 for *s*-stem adjectives in \*-ēs. A heteroclitic adjective can be reconstructed as well: \*piHwō(n) ‘fat’ (Ved. *pívan-*, Gr. πῖων) – f. \*piHwerih<sub>2</sub> (Ved. *pívarī*, Gr. πῖερα). As can be seen, some adjectives had separate masculine/feminine/neuter forms (e.g., all thematic forms), while others had identical masculine/feminine forms (as in some nominal stems). There were a number of adjectival suffixes, cf. \*med<sup>w</sup>-y-os (p. 31), \*h<sub>1</sub>rud<sup>h</sup>-r-os (p. 19), \*-went- ‘having . . .’ (Ved. *putrávant-* ‘having a son’ from *putrás* ‘son’), etc. The **Caland System** is a name for a pattern of adjectives in \*-ro-, \*-mo-, \*-u-, \*-nt- (often occurring with the same root, cf. Gr. ἐλαφ-ρ-ός ‘light (in weight), nimble, small’, ἐλαχ-ύ-ς ‘small, short’) that have \*-i- instead as the first element in compounds (cf. Gr. κῦδ-ρ-ός ‘glorious’ but κῦδ-ι-άνειρα ‘bringing men glory’) and that often go together with neuter *s*-stem nouns (Gr. κῦδος ‘glory’ ⇒ Eng. *kudos*) and stative \*-eh<sub>1</sub>- verbs (cf. Gr. ἐρυθ-ρ-ός, OCS *rub-d-rъ* ‘red’, and Lat. *rub-ē-re* ‘to be red’) formed from the same root.

A couple of adjectival/adverbial prefixes appearing in compounds can be reconstructed: \*h<sub>3</sub>- ‘un-’ (Ved. *a-*, Gr. *á-*, Lat. *in-*, Goth. *un-*, OIr. *an-*; zero grade of the particle \*ne ‘no’ – p. 54; cf. the famous PIE collocation \*Klewos h<sub>3</sub>-d<sup>h</sup>g<sup>w</sup>hitom ‘imperishable fame’ > Ved. *śrávo* . . . *ákṣitam*, Gr. Hom. κλέος ἄφθιτον), \*h<sub>1</sub>su- ‘good’ (p. 47,

cf. \*h<sub>1</sub>sumenēs – p. 79), \*dus- ‘bad’ (Ved. *duṣ-*, Gr. *δυσ-*, Goth. *tuz-*, cf. Gr. *δυσ-μενής* ‘hostile’, OCS *dъždь* ‘rain’ < \*‘bad day’), \*sēmi- ‘half’ (Gr. *ἡμι-*, Lat. *sēmi-*, OEng. *sām-*), etc.

### Comparison of adjectives

PIE did not have the full-fledged paradigmatic comparison that later and modern IE languages have. It apparently had a number of derivation suffixes (similar to Eng. *-ish* in *smallish*, etc.) that modified the meaning of adjectives, some of which became regular comparative/superlative suffixes in later languages. The most widespread comparative(-like) suffix was \*-yōs (m./f.) (Lat. *senior* |v|, arch. Lat. *-ō* |v|, OIr. *sinu* ‘older’) – \*-yos (n) (Lat. *melius*, OCS *bolje* ‘better’, perhaps also \*-is, cf. adverbial Lat. *magis* ‘more’, Goth. *mins* ‘less’, OPruss. *tālis* ‘farther’) – oblique \*-is- (OCS oblique *dražь* |š| ‘dearer’, Goth. *batiz* |a| ‘better’), and probably \*-yes- in some cases (Ved. oblique *-yas-*), with +±± holodynamic paradigm and a full root even from adjectives with a zero root in the positive, cf. Gr. (Ion.) *κρέσσων* |v| ‘stronger’ < \*kret-yōs from *κρατός* ‘strong’ < \*k<sub>ṛ</sub>t-u-s; Ved. *drāgh* |ī|yas- ‘very long; longer’ < \*dleHg<sup>h</sup>-yos- from *dīrghás* ‘long’ < \*d<sub>l</sub>Hg<sup>h</sup>-os. The zero-grade \*-is- also appeared in complex superlative suffixes \*-is-t(H)o- (Ved. *náviṣṭhas* ‘newest’, Gr. *βέλτιστος*, Goth. *batists* ‘best’, made from a full-grade root and not from the positive, cf. Ved. *téj-iṣṭh-a-* from *tig-m-á-* ‘sharp’), pairing with the comp. \*-yos-, and dialectal/innovative \*-is-ṛ-Ho- (Lat. *māximus* ‘biggest’, OIr. *nessam* ‘nearest’). The suffix \*-ter-o-, originally used in contrastive (positive) meaning in adjectives (Lat. *dexter*, Gr. *δεξιτερος* ‘right’, Ved. *úttaras* ‘upper’) and pronouns (\*k<sup>w</sup>oteros ‘which (of two)’ > Ved. *katarás*, Gr. *πότερος*, OCS *kot* |o|ryi ‘which’, Goth. *hvapar*), later yielded the comparative in Greek (*σοφώτερος* ‘wiser’) and Indo-Iranian (Skr. *mahattaras* ‘greater’). The complex suffix \*-ṛṇ-H-o- often yielded a superlative (Ved. *utámás* ‘highest’, Lat. *optimus* ‘best’, Gaul. name *Uertamocorii* ‘highest warriors’; in Greek \*-ṛṇ-t-o-, cf. *σοφώτατος* ‘wisest’), pairing with the comp. \*-tero-. It is possible that in some cases an unmarked PIE adjective might have functioned as a comparative/superlative, cf. Hitt. *šallayaš šiunaš šalliš* ‘the greatest (lit. great) of the great gods’, and p. 438 for the Armenian comparative. An interesting but not completely clear phenomenon in IE languages is that for frequent/basic adjectives (‘good’, ‘bad’, ‘small’, ‘big’) one often finds suppletive forms, which usually differ from language to language, cf. for ‘good’ – ‘better’ – ‘best’ Gr. *ἀγαθός* – *βελτίων* – *βέλτιστος* (with other variants), Lat. *bonus* – *melior* – *optimus*, OCS *dobrъ* – *bolii* (*lučii*), Goth. *goþs* – *batiza* – *batists*, OIr. *maith* – *ferr* – *deg*. Cognates among the mentioned suppletive adjectives are just \*bel-/mel- (cf. p. 17) ‘better’ and \*min- ‘smaller’ (Lat. *minor*, OCS *мънii*, Goth. *minniza*).

### ADVERBS

PIE did not have general, special adverbial suffixes for derivation from nouns and adjectives (cf. Eng. *man-ly*, *nice-ly*). Instead, various (usually singular) nominal and adjectival cases (all except the vocative) were used adverbially. However, not a lot can be reconstructed directly for PIE since this type of adverb formation was very productive, and the forms are widely diverse in IE languages; cf., e.g., the adverbial ‘by night’ expressed by the gen. sg. in Gr. *νυκτός*, Goth. *nahts*, the acc. sg. in Ved. *náktam*, and the instr. sg. in OCS *noštъjъ* ‘by night’. Neuter forms (in the nom./acc. sg.) are generally often used adverbially, cf. the neuter adjectival \*megh<sub>2</sub> (p. 79) used as an adverb (Ved. *máhi* ‘greatly’, Gr. *μέγα* ‘very much’, ONor. *mjōk* ‘much, very’, Hitt. *mekk* |i| ‘very, greatly’), or Ved. *purú* ‘abundantly’, Gr. *πολύ* ‘much’, Goth. *filu* ‘much, very’; OCS *много* ‘much’,

etc. For the adverbial use of the locative, cf. Lat. *temere* ‘blindly’ (< \*‘in the dark’), OCS *dolě* ‘below’, *dobrě* ‘good’, etc.; for the instrumental, cf. Ved. *dīvā* ‘by day’, *madhyā* ‘in the middle’, Gr. κρυφή ‘in secret’ (p. 66), OCS *vbčera* ‘yesterday’ (p. 64), etc.

Some apparently non-derived (or not easily analyzable) adverbs can also be reconstructed, e.g., \*d<sup>h</sup>yes (p. 53) ‘yesterday’ (Ved. *hyás*, Gr. χθές, Lat. *her|ī*, OIr. *|in|-dé*, Alb. *dje*), originally perhaps a loc. sg.; \*nū(n) ‘now’ (Ved. *nū*, Gr. νῦν, Lat. *nun|c|*, OCS *ny|ně* and CS *nb|ně*), Goth./OIr. *nu*, Hitt. *nu* ‘and, but’, *|ki|nun* ‘now’, etc. – p. 54–55), usually connected to the adj. \*new-o-s ‘new’, etc. The later verbal augment \*h<sub>1</sub>e (p. 6, 95, 161) was originally probably an adverb. A few postpositions (p. 104–105) were originally probably adverbs or both postpositions and adverbs, cf. \*(s)upo ‘below; under’ (Ved. *úpa* ‘up to, upon, above, near’, Gr. ὑπό/ὑπο ‘(from) under’, Lat. *sub* ‘under, below, up to’, Goth. *uf* ‘under’, OIr. *fo* ‘under’) and \*(s)uper(i) ‘above; over’ (Ved. *upári* ‘above’, Gr. ὑπέρ/ὑπερ ‘over, above’, Lat. *super* ‘over, above’, Goth. *ufar* ‘over, above’, OIr. *for* ‘over’). Some adverbial suffixes did exist; for \*-d<sup>h</sup>e, \*-r, and pronominal adverbs generally, cf. p. 89. For numerical adverbs, cf. p. 91; for adverbial compound prefixes like \*h<sub>1</sub>su- ‘good’, cf. p. 79–80.

## PRONOUNS

Unlike nouns/adjectives and verbs, which always had CVC type roots, pronominal roots were usually CV, VC, or V (of course, \*HV- is always theoretically possible for initial \*V-, cf. p. 52). The pronominal declension had mostly the same endings as the nominal one, except for five special different endings (apparently, not all appeared in all pronominal declensions): nom. sg. m. \*-Ø (cf. nominal \*-s), nom. sg. n. \*-d (cf. nominal \*-Ø/m), gen. sg. m./n. \*-(e)so (cf. nominal *o*-stem \*-osyo, otherwise \*-es/-os), nom. pl. m. \*-i (cf. nominal \*-es), and gen. pl. \*-som (cf. nominal \*-om, p. 63). The oblique m./n. sg. cases (gen./abl./dat./loc.) of non-personal pronouns also feature the element \*-sm- before the endings, usually thought to be the zero grade of \*som- ‘one’ (p. 89); cf. the English use of *this one*, *which one*, etc. In the plural (and dual) masculine/neuter forms, the nom. \*-i often spreads already in PIE to the oblique cases as part of the enlarged stem (e.g., gen. pl. \*to-y-som, p. 86). Many basic pronouns had suppletive forms (cf. heteroclitic nouns – p. 77–78). Personal pronouns, unlike other pronouns and nouns, did not have all seven/eight cases (this is opposite to the situation in many present-day IE languages – like English, the Romance languages, and Macedonian/Bulgarian – where personal pronouns preserve the declension, while the nouns do not). In the history of the IE languages, pronominal and nominal forms often mixed, cf. \*-o-i instead of original \*-o-es in the nom. pl. *o*-stems in Latin/Greek/Balto-Slavic (p. 65), or *-rum* instead of the older *-um* in the gen. pl. in Latin (p. 67). Non-personal pronouns, which have different gender forms, like adjectives, sometimes had two (m. = f., n.), sometimes three forms (m., f., n.).

### Personal pronouns

The personal pronouns are rather difficult to reconstruct since the attested IE systems are in many cases rather different (due to copious innovation).<sup>7</sup> They were characterized by partial (nom. sg. \*tu ‘you’ – gen. sg. \*teb<sup>h</sup>e) or complete suppletivity (nom. sg. \*h<sub>1</sub>eǵ ‘I’ – acc. sg. \*me); by the existence of clitic forms – sometimes as the only existing, or at least reconstructable, form and often syncretic (like gen./dat./acc. pl. \*nos ‘us’); and apparently by the lack of reconstructable loc./instr. cases (all languages have different, innovative-looking forms), which were either non-existent or, less likely, innovative in all languages – the lack of separate vocative forms is not strange, and the ablative is the same as the genitive

in most nominal declensions anyway. Like many older IE languages, PIE had no separate 3rd person forms of personal pronouns – demonstrative/anaphoric pronouns (p. 85) were used instead. Many monosyllabic forms exhibited facultative lengthening (p. 54–55). The reconstructed forms are not always completely morphologically transparent (i.e., the root, suffixes, and endings are not always easily distinguished – e.g., in gen. \*mene ‘of me’).

### 1 sg. *T'*

- nom. \*h<sub>1</sub>eǵ > GAv. *as-cī̄t* (1x), Slav. \*jǎ (e.g., BCMS *jā*), OLith. *e|š|*, Hitt. *|ū|k*  
 \*h<sub>1</sub>eǵ-Hom > Ved. *ahám*, GAv. *azəm*, Gr. *ἐγώ*, Slav. \*jazь (e.g., OCS *azъ*),  
 Goth. *ik*, Arm. *e|s|*  
 \*h<sub>1</sub>eǵ-oh<sub>2</sub> > Gr. *ἐγώ*, Lat. *egō*, Ven. *ego*
- acc. \*mē (accented/clitic) > Ved. *mā*, GAv. *mā*, Gr. *|ē|μῆ*, clitic *μῆ*, Lat. *mē*, OCS *m|ē|*,  
 OPruss. *mie|n|*, Goth. *mi|k|*, OIr. *mé*
- gen. \*mene > Ved. *má|ma*, YAv. *mana*, OCS *mene*  
 \*moy (clitic) > Ved. *me*, GAv. *mōi*, Lat. *m|s|*
- dat. \*me-ǵ<sup>h</sup>-i > Ved. *máhy|a(m)|*, Lat. *mih|ī|* (Umbr. *meh|e|*), Arm. *|i|nj*  
 \*moy (clitic) > Ved. *me*, GAv. *mōi*, Gr. *μοι*, OCS *mi*, OLith. *mi*

The 1 sg. pronoun showed the suppletivity of \*h<sub>1</sub>eǵ- (nom.) and \*m- (oblique). The oblique cases had both stressed and clitic forms. Three nom. forms are reconstructable, one non-derived (\*h<sub>1</sub>eǵ) and two with suffixes, the \*-oh<sub>2</sub> one perhaps being dialectal. The existence of two (even three) parallel forms may not be as strange as it might seem (even without any specific different function/meaning, which may have existed in PIE), cf. Hitt. *ūk* and *ukel*, *ukila* ‘I, myself’. Hitt. oblique *amm-* (cf. acc. *ammuk* ‘me’), Arm. *i-* (cf. acc. *is* ‘me’), and Gr. *ἐ-* in stressed forms (cf. acc. *ἐμέ* but clitic *μῆ*) would point to \*h<sub>1</sub>m- (p. 48), not \*m-. However, this is not in accord with the glottogonic, but nonetheless persuasive, connection of pronominal \*me- with verbal eventive 1 sg. \*-m, 1 pl. \*-me (p. 93). Thus, it is perhaps better to assume that this \*h<sub>1</sub>m- is innovative – an analogy to the initial \*h<sub>1</sub>- in the nom. (where \*h<sub>1</sub>eǵ, not \*eǵ, has to be reconstructed in that case) or later to the initial \*-e- in the nom. However, \*h<sub>1</sub>m- would formally also work for PIE. Like in the 2nd person accented dat. sg., the ending is difficult to reconstruct (all languages show different endings). Here, we tentatively reconstruct PIE \*-i, on the basis of Ved. *-yam*, OLith. dat. *mani*, and perhaps clitic \*mo-i, but that is admittedly highly speculative. The element \*-ǵ<sup>h</sup>- in the dat. is usually thought to be some kind of particle originally. It is possible that \*moy had a variant \*mey (which would be a more ready pre-form for Latin, Lithuanian, and Old Church Slavic). In the ablative, Ved. *mád* and OLat. *mēd* are formal cognates (PIE \*me-t?) but are more likely separate innovations.

### 2 sg. ‘you’

- nom. \*tū > Ved. *t(u)|vám|*, GAv. *tū*, Gr. (Dor./Aeol.) *τῷ*, Lat. *tū*, OCS *ty*, Lith. *tù* (OPruss. *tū*), Goth. *þu*, OIr. *tú* (also *tu-ssu*), Arm. *|d|u*, Alb. *ti*, Toch. B *t(u)|we|*
- acc. \*twē > Ved. *tvā|m|*, GAv. *ǰβā*, Gr. *σέ*  
 \*tē (clitic) > Gr. (Dor.) *τέ*, Lat. *l|ē|*, OCS *t|ē|*, OPruss. *tie|n|*, Goth. *þi|k|*
- gen. \*tewe > Ved. *táva*, OCS *te|v|e*, Lith. *tav-*, Toch. B *ci*  
 \*toy (clitic) > Ved. *te*, GAv. *tōi*, Lat. *t|s|*
- dat. \*tub<sup>h</sup>-i > Ved. *túbhy|a(m)|*; \*teb<sup>h</sup>-i > Av. *ta|bi|ā|*, Lat. *tib|ī|* (Umbr. *tef|e|*), OCS *teb|ē|*, OPruss. *tebb|e(i)|*  
 \*toy (clitic) > Ved. *te*, Gr. (Dor./Lesb./Ion.) *τοι*, OCS *ti*, OLith. *ti*

The forms are in many ways parallel to those of the 1 sg. The stems \**tew-/tu-/tw-* and \**t-* are marginally suppletive (these are probably ultimately somehow related to the verbal 2 pl. \**-te*). Anatolian languages show \**ti-* in the nom. (CLuw. *tī*) and \**tu-* in oblique cases (Hitt. *tu-*), which is difficult to explain as innovative so some think that PIE originally had a nom. \**tī* (cf. Kloekhorst 2008: 112–115), though the supposed original Anat. nom. \**tu-* is often thought to be the origin of the vocalism of Hitt. *tūk* ‘I’. The existence of both short and long reflexes in the nom. points to monosyllabic lengthening (p. 54–55), as in other similar personal pronouns forms (cf. acc. \**mē* and \**t(w)ē*), not to a laryngeal length. Ved. *túbhya(m)* with *tu-* looks like an archaism (cf. OCS *mъně* ‘to me’, Latv. dial. *mun-* by possible early analogy), while other languages point to a more regular \**teb<sup>h</sup>-* (cf. gen. \**tewe* and dat. \**meǵ<sup>hi</sup>*). The usual dat. ending \**-ey* in Lat. *tibī*, OPruss. *tebbe(i)* is perhaps an innovation (by analogy to nouns). Unlike the 1 sg. pronoun, the stressed and unstressed forms were distinguished in the acc.

### 1 pl. ‘we’

- nom. \**we-y(-es)* > Ved. *vay|ám*], GAv. *va|ēm*], Goth. *wei|s*], Hitt. *wē|š*], Toch. B *we{s}* (also archaic \**mes?*, see below)
- acc. \**ŋ-s-me* > Ved. *asm|ā|n*], YAv. *ahma* (GAv. *āhmā*), Gr. (Aeol.) *ἄμμε*, Goth. *uns*, Hitt. *anz|ā|š*]  
\**nō-s* (clitic) > Ved. *nas*, GAv. *nā* (< \**nās*), Lat. *nōs*, OCS *n|y*], Hitt. *-naš*, Alb. *ne/na*
- gen. \**nō-s* (clitic) > Ved. *nas*, GAv. *nā* (< \**nas*), Lat. *nos|trum*], OCS *nas|b*], OPruss. *n|oū|s|an*], Alb. *na*
- dat. \**nō-s* (clitic) > Ved. *nas*, GAv. *nā* (< \**nas*), OCS *n|y*], Hitt. *-naš*, Alb. *na*

The suppletive forms were \**we-* and \**no-/ŋ-* with plural endings (the pronominal \**-i* in the nom. and \**-s(-)*, presumably a variant of the nominal pl. \**-es*, in other cases). The form \**wey* is not really attested as such, since Indo-Iranian has the usual additional \**-om* and Germanic/Anatolian point to \**wey-es* with the additional plural \**-es*. For \**we-* cf. the dual pronominal and verbal forms (p. 94), while \**ŋs-* is often derived from \**ŋs-* (cf. the nominal acc. pl. – p. 63), i.e., from \**me-* as in the 1 sg. pronoun (cf. also verbal 1 sg. \**-m*) and \**me-* in 1 pl. verbal forms. Because of these considerations, one is prone to consider \**mes* ‘we’ (Lith. *mēs*, Arm. *mek*) an archaic dialectal by-form (with \**we-* often considered originally dual), though it could also be an innovation due to analogy to the verbal ending \**-mes* (p. 93). The acc. \**(ŋs)me-* is an unlikely source for \**mes*. For the gen./dat., only clitic forms seem to be reconstructable. In the stressed acc., a suffix (or reduplication of the original root?) \**-me* can be reconstructed on the basis of Greek and Avestan. It is not impossible that the dat. had a non-clitic \**ŋs-m-* with an unreconstructable ending (cf. Ved. *asm-é*, Av. *ahm-a’biū*, Gr. Aeol. *ἄμμ-ι(v)*, Goth. *uns-(is)*), but this can also be a later analogy to the acc. (apparently, this \**-sm-* had nothing to do with \**-sm-* in non-personal pronouns, see below).

### 2 pl. ‘you’

- nom. \**yū-s* > Ved. *yū|yám*], GAv. *yūš*, OCS *v|y*, Lith. *jūs*, Goth. *jus*
- acc. \**u-s-we* > Goth. *l|zwi|s*], MW *l|chwī*]; *us-me* > Ved. *l|y|ušm|ā|n*], Gr. (Aeol.) *ὑμμε*, Hitt. *šumm|ā|š*]  
\**wō-s* (clitic) > Ved. *vas*, GAv. *vā* (< \**vās*), Lat. *uōs*, OCS *v|y*], OPruss. *wa|n|s*
- gen. \**wō-s* (clitic) > Ved. *vas*, GAv. *vā* (< \**vas*), Lat. *uos|trum*], OCS *vas|b*]
- dat. \**wō-s* (clitic) > Ved. *vas*, GAv. *vā* (< \**vas*), OCS *v|y*]



The forms are in many ways parallel to those of the 1 pl. The suppletive forms are \*yu- (nom.) and \*wo-/u- (oblique). A frequent *ad hoc* assumption is that \*yu- stems from the older \*u-. The nom. (which, unlike the 1 pl., has nominal \*-s) is attested with a long \*ū only, so a laryngeal \*yuHs is theoretically possible instead of the monosyllabic lengthening (p. 54–55), but not likely (cf. the dual form). The stressed acc. again shows \*-me, but Goth. *izwis* looks archaic if we consider this \*-we (also seen in the dual) an original reduplication of the root. Cf. Katz 1998 for \*-me/-we.

### 1 du. ‘us two’

nom.	*wě > Ved. <i>vā m</i> (1x), GAv. <i>vā</i> , OCS <i>vě</i> , OLith. <i>ve du</i> , Goth. <i>wi t</i>
acc.	*ṅh <sub>1</sub> -we > Ved. <i>āv{ām}</i> , GAv. <i>āḍāuuā</i> , Goth. <i>ug{kis}</i>
	*no-h <sub>1</sub> (clitic) > Ved. <i>nā u</i> , Av. <i>nā</i> , Gr. (nom./acc.) <i>ωό</i> , OCS <i>na</i>
gen./dat.	*no-h <sub>1</sub> (clitic)

The suppletive forms are much the same as in the 1 pl. Cf. also the verbal 1 du. \*-we- (p. 94). We reconstruct the nom. form with monosyllabic lengthening because of the short and long reflexes. If one were to reconstruct \*weh<sub>1</sub>, OLith. *ve-du* and Goth. *wi-t* would be difficult to account for. The nom. form is the same as the plural one, except for the missing plural \*-i. The laryngeal length of the gen./dat./acc. clitic (with the dual ending \*-h<sub>1</sub>) is proven by consistent long reflexes, parallelism to the plural forms (which are the same except for having the plural ending \*-s(-)), and the zero-grade length in the stressed acc. form (\*ṅh<sub>1</sub>-). The clitic was probably \*nōh<sub>1</sub> (like pl. \*nōs), but no language distinguishes \*ōH from \*oH. The reconstruction of the gen./dat. clitic is due to the supposed structural parallelism to the plural forms. In dual forms, the stressed acc. has only \*-we (unlike the prevalent \*-me in the plural), according to the reflexes. Some think this \*-we is found also in the acc. sg. \*t-we (but this may also be \*tu-e, cf. nom. \*tu).

### 2 du. ‘you two’

nom.	*yū > Ved. <i>yū vām</i> , OCS <i>ʋy</i> , Lith. <i>jū du</i>
acc.	*uh <sub>1</sub> -we > Ved. <i>yū v{ām}</i>
	*wo-h <sub>1</sub> (clitic) > Ved. <i>vā m</i> , OCS <i>va</i> , Toch. B nom. <i>we ne</i>
gen./dat.	*wo-h <sub>1</sub> (clitic)

The forms are parallel to the 1 du., and the suppletive stems to ones in the 2 pl. forms.

## Reflexive pronouns

The reflexive pronoun ‘self’ was identical to the 2 sg. personal pronoun, except that it did not have a nom. sg. and had an initial \*-s- instead of \*t-. Cf. gen. \*sewe (OCS *seʋe*), dat. \*seb<sup>hi</sup> (YAv. *h|uu|āuuōii|a|/hwawya/*, OLat. *sib{ei}*, OCS *seb{ě}*, OPruss. *sebb{ei}*), clitic gen. \*soy (GAv. *hōi*), clitic dat. \*soy (GAv. *hōi*, Gr. *oi*, OCS *si*), acc. \*swē (Ved. *svalyām*, Gr. *ἑ*, P amph. *ḗē*), clitic \*sē (Lat. *sē*, OCS *sʋē*, OPruss. *sie|n*, Goth. *si|k*). The pronoun was originally used not only for the 3 sg. but for the 1 and 2 sg. as well.

## Possessive pronouns

The singular possessive pronouns, derived from personal and reflexive pronouns, can be formally reconstructed, though their antiquity is not beyond doubt (the forms often differ

in specific languages): \*m-o-s (Gr. |ἐ|μός, GAv. *mā*) and \*me/oy-o-s ‘my’ (Lat. *meus*, OCS *mojъ*), \*tw-o-s ‘thy’ (Ved. *tvás*, GAv. *ββ̄ā*, Gr. σός, Arm. *k’o*), \*sw-o-s ‘own’ (Ved. *svás*, Gr. ὄς).

**Demonstrative pronouns**

A number of demonstrative pronominal stems can be reconstructed – some better attested than others. Considering the number of the attested stems, it is probable that PIE had a three-way system of demonstrative pronouns (like Lat. *hic/iste/ille* and most IE languages, or Japanese *kore/sore/are*), correlated with three persons, and not a two-way one (like English *this/that*), cf. p. 160. The exact reconstruction of the semantics of the stems is a bit speculative, since these change easily in languages. The three best-attested (and probably main) demonstratives are \*ki- ‘this (near me/speaker)’, \*s-/t- ‘that (near you/the addressed)’, and \*(h<sub>1</sub>)ey-/(h<sub>1</sub>)e- ‘that (near him/over there)’. The somewhat less well-attested stems are \*(h<sub>1</sub>)e/ow-o-s ‘this’ (Ved. *avás*, OCS *ovъ*) and \*(h<sub>2/3</sub>)e/on- ‘that (near him)’ (OCS *onъ*, Lith. *anàs*, Arm. *na*). It is not clear why there were five different reconstructable demonstrative stems (a number of scenarios are possible), though some languages preserve all of them, cf. OCS *sb*, *ovъ*, *tъ*, *onъ* (suppletive acc. *i*).

For the pronoun \*ki- ‘this’ (also \*ko-?), cf. OCS *sb* (m.) – *si* (f.) – *se* (n.), Lith. *šis* – *ši* (f.), OIr. *cé*, Arm. *sa*, Hitt. *kāš* (CLuw. *zāš*) – *kī* (n.), PGmc \**xī* (OEng. *hē* ‘he’, Goth. *himma daga* ‘today’, etc.), also Lat. (*hi*)-*c* (cf. *cis* ‘on this side’), Gr. (ἐ)κεῖ-voς ‘that one there’ (\*key-?), Alb. *si-vjet* ‘this year’. Owing to differences in languages and partial attestations, the PIE declension is not clear.

The suppletive forms are well attested in \*s-o (m.) – \*s-eh<sub>2</sub> (f.) – \*t-o-d (n.) ‘that’ (cf. Av. *hō* – *hā* – *taṭ*, Gr. ὁ – ἡ – τὸ ‘the’, Cib. *so* (?) – *sa* – *soḍ*, Toch. B *se* – *sā* – *te*, OLat. f. *sa-psa* ‘herself’). The \*s- appears only in the nom. sg. m./f.; all other forms have a \*t-. Here is the declension (n. = m., except in the nom./acc. and du.):

**TABLE 1.29 IE DEMONSTRATIVE PRONOUN \*SO – \*SEH<sub>2</sub> – \*TOD (SINGULAR)**

sg.	PIE	Ved.	OCS	Lith.	Goth.
nom.	*s-o – *s-eh <sub>2</sub> – *t-o-d <sup>a</sup>	<i>sá( s ) – sá – tád</i>	<i>ʌṭ – ʌʌ – to</i>	<i>ʌʌ s  – ʌʌ<sup>b</sup></i>	<i>sa – so – þat a </i>
acc.	*t-o-m – *t-e-h <sub>2</sub> -m – *t-o-d	<i>tám – táṃ – tád</i>	<i>tъ – tḡ – to</i>	<i>tā</i>	<i>þan a  – þo – þat a </i>
gen.	*t-e-so – *t-e-sy- e-h <sub>2</sub> -es	<i>tás y a – tásyās</i>	<i>tʌ{g}o – tʌojə\</i>	<i>t{ð} – t{ðs}</i>	<i>þis – þiʌʌos</i>
abl.	*t-o-sm-ōt – *t-e-sy- e-h <sub>2</sub> -es	<i>tásmāt – tásyās</i>			
dat.	*t-o-sm-ōy – *t-o-sy- e-h <sub>2</sub> -ey	<i>tásmai – tásyai</i>	<i>tomu – toʌ\</i>	<i>tamui<sup>c</sup> – t{aī}</i>	<i>þamma – þiʌʌai</i>
loc.	*t-o-sm-i – *t-o-sy- e-h <sub>2</sub> -i	<i>tásmi n  – tásy{ām}</i>	<i>tomъ – toʌ\</i>	<i>tam{è} – toʌ\{è}</i>	
instr.	*t-o-y-? – *t-o-y- eh <sub>1</sub> (?)	<i>ténā – táyā</i>	<i>tēmъ – toʌ\φ\</i>	<i>t{uō} – t{q\}</i>	

<sup>a</sup> In sg./pl. always m. – f. – n. and m./n. – f. (when just two forms); <sup>b</sup> Lithuanian has no neuter, cf. OPruss. *sta* < \*s-t-od; <sup>c</sup> Old Lithuanian.

The \*-o/-e- that appears after \*s-/t- in the sg./pl. is apparently a thematic vowel. The non-pronominal endings are from o-/eh<sub>2</sub>-stems (except for masculine loc. \*-i, not \*-oy).

Ved. *sá*, Gr. *ó*, and Goth. *sa*, with no ending, are archaic (cf. also \**ey* below), though the Greek declension of the pronoun is otherwise innovative, with all nominal endings. Goth. *þat-a* probably proves the neuter ending was \*-d (not \*-t). Germanic (and Slav. *česo*, Gr. Ion. *τέο*) probably points to a gen. in \*-e-so, replaced by thematic \*-o- and other endings elsewhere. In the masculine abl./dat./loc., the element \*-sm- (Ved. *-sm-*, OCS/Lith. *-m-* with a special development of \*-sm-, OPruss. *-sm-*, Goth. *-mm-*) appears after the root, cf. also OPruss. dat. *stesmu* (with \*s- + \*t-), Clb. dat. *somui* (with generalized \*s-), and Arm. dat. *dma* < \*doma (with *d-* as in *du* ‘thou’, p. 438). The f. gen./abl./dat./loc. sg. forms appear, on the basis of Indo-Iranian, to have \*-sy- instead (connected to \*si- ‘one’ > Hitt. *šī-*, Gr. Hom. *ῥα*, besides *μία*, cf. Beekes 1988: 81; or perhaps from older \*-smy- from \*smih<sub>2</sub> ‘one’ – p. 89). In Old Church Slavic, \*tosy- is replaced with \*toy- (as in the pl. and instr. sg.); in Gothic with \*tes- (as in the masculine gen.); in Lithuanian (and Greek) with a simple \*t- plus endings (e.g., the gen. \*t-eh<sub>2</sub>es instead of \*tesy-eh<sub>2</sub>es). The instr. forms are unclear, though both Vedic and Old Church Slavic point to \*toy- (otherwise the pl. stem).

TABLE 1.30 IE DEMONSTRATIVE PRONOUN \*SO – \*SEH<sub>2</sub> – \*TOD (PLURAL)

pl.	PIE	Ved.	OCS	Lith.	Goth.
nom.	*t-o-y – *t-e-h <sub>2</sub> -es – *t-e-h <sub>2</sub>	<i>té – táś – tá</i>	<i>ti – tʲ\ – ta</i>	<i>tiẽ – tōś</i>	<i>þai – þos – þo</i>
acc.	*t-o-ns – *t-e-h <sub>2</sub> (n)s – *t-e-h <sub>2</sub>	<i>táñ – táś – tá</i>	<i>ty (m./f.) – ta</i>	<i>tuōś – tās<sup>d</sup></i>	<i>þans – þos – þo</i>
gen.	*t-o-y-som – *t-e-h <sub>2</sub> -som	<i>téśāñm – táśāñm</i>	<i>těxъ</i>	<i>tāŋ\<sup>e</sup></i>	<i>þʲāz{e} – þʲāzō<sup>g</sup></i>
dat.	*t-o-y-b <sup>b/</sup> mos – *t-e-h <sub>2</sub> -b <sup>b/</sup> mos	<i>tébhʲyas – tábhʲyas</i>	<i>těmъ</i>	<i>tiems – tóms<sup>f</sup></i>	<i>þaim</i>
loc.	*t-o-y-su – *t-e-h <sub>2</sub> -su	<i>téśu – táśu</i>	<i>těxъ</i>	<i>tuos{è} – tos{è}</i>	
instr.	*t-o-y-b <sup>b/</sup> i(s) – *t-e-h <sub>2</sub> -b <sup>b/</sup> i(s) <sup>a</sup>	<i>tébhis<sup>b-</sup> – tábhis</i>	<i>těmi<sup>c</sup></i>	<i>t{aĩs} – tomis</i>	

<sup>a</sup> BSL. \*-miHs; <sup>b</sup> AV *táś* (cf. Av. *tāiš* and Lith. *taĩs*) is innovative; <sup>c</sup> Old Church Slavic (and Gothic in the dat.) have generalized \*toy- in both m./f. in the gen./dat./loc./instr. (cf. the Lat. f. gen. *is{tārum}*); <sup>d</sup> cf. OPruss. *stans*; <sup>e</sup> cf. OPruss. *steison*; <sup>f</sup> OLith. *-mus*; <sup>g</sup> cf. adjectival *-aize*, *-aizo*.

The nom. and gen. have pronominal endings. From forms like \*toy, many later languages generalized the originally pronominal ending \*-oy in the nominal *o*-stems as well (p. 64–65). For gen. \*-som (often changed to \*-sōm, as in the nominal \*-om/-ōm, p. 70), cf. also Clb. *soisum* and Hitt. *kinzan* (from \*Ki-). This \*toy- seems to have spread from the nom. to the gen./dat./loc./instr. pl. (and instr. sg.).

TABLE 1.31 IE DEMONSTRATIVE PRONOUN \*SO – \*SEH<sub>2</sub> – \*TOD (DUAL)

du.	PIE	Ved.	Gr.	OCS
nom./acc./voc.	*t-o-h <sub>1</sub> (m.) – *t-o-yh <sub>1</sub> (f./n.)	<i>tā (m.) – té (f./n.)</i>	<i>τὸ</i>	<i>ta (m.) – tē (f./n.)</i>
gen./loc.	*t-o-y-ow	<i>tāyo s </i>	<i>τοῖ{v}</i>	<i>toju</i>
dat./abl./instr.	*t-o-y-b <sup>b/</sup> y/moh <sub>1</sub>	<i>tābhyā m </i>	<i>τοῖ{v}</i>	<i>těma</i>

The dual forms have the nominal endings, with the oblique \*-oy- as in the plural.  
 The pronoun for ‘that (over there)’ (used anaphorically as well – referring to something already mentioned – which was perhaps its original function) has suppletive \*(h<sub>1</sub>)ey- (zero-grade \*(h<sub>1</sub>)i-) and \*(h<sub>1</sub>)e- (n.=m., except in nom./acc.):

TABLE 1.32 IE DEMONSTRATIVE PRONOUN \*EY (\*IS) – \*IH<sub>2</sub> – \*ID (SINGULAR)

sg.	PIE	Ved.	Lat.	OCS	Goth.
nom.	*ey (*i-s) – *i-h <sub>2</sub> – *i-d	ay ám  – i yám  – id ám  <sup>a</sup>	is <sup>b</sup> – {ea} – id <sup>c</sup>	{onъ – ona – ono i <sup>g</sup> }	is –  s i – it a  <sup>h</sup>
acc.	*i-m – *i-h <sub>2</sub> -m – *i-d	im ám  – {imám  – id ám	OLat. im – {eam} – id	i <sup>g</sup> – {jǫ – je}	in a  – i j a  – it a
gen.	*e-so – *e-sy-e-h <sub>2</sub> -es	as y á – asyás	e i u s  <sup>d</sup>	je{g}o – je{j}e\	is <sup>i</sup> – iʒʌos
abl.	*e-sm-öt – *e-sy-e-h <sub>2</sub> -es	asmát – asyás	e{ō} – e{ā}		
dat.	*e-sm-ōy – *e-sy-e-h <sub>2</sub> -ey	asmái – asyái	e{ī}	jemu – jeʌ\	imma – iʒʌai
loc.	*e-sm-i – *e-sy-e-h <sub>2</sub> -i	asmī n  – asy{ám}	<sup>e</sup>	jemb – jeʌ\	
instr.	*ey-? – (?)	ená – ayá		imъ – jeǰo	

<sup>a</sup> ‘this here’; <sup>b</sup> OLat. e|s| (1x); <sup>c</sup> ‘that one; he/she/it’; <sup>d</sup> from \*e-syo-s; <sup>e</sup> cf. South Picene loc. **esme/in**; <sup>f</sup> ‘he, she, it’, cf. above for *on-*; the original forms are seen in the determinate adjective endings -i < \*jъ, {-ja}, {-je}; <sup>g</sup> < \*jъ; <sup>h</sup> ‘he/she/it’ (cf. Germ. *er – sie – es*); <sup>i</sup> OHG *es* (n.).

Many forms are parallel to those of \*s-/t- (cf. above), but the nature of the suppletion pattern is different and the reconstruction more difficult and less secure in some cases. The variant \*ey-/i- appears in the nom./acc. sg. and \*e- elsewhere. For the nom. cf. also GAv. m. *aii|ām|* ‘this’ – f. *i|m|* ‘this’, n. *iṭ* ‘it’; Lith. *j|is* ‘he’ – *j|i* ‘she’ (this pronoun mixes with the relative \*yo- in Balto-Slavic), and OIr. *é* ‘he’ – *|s|i* ‘she’ – *ed* ‘it’. The archaic nom. m. form was perhaps, though a bit speculatively, \*ey-Ø with a pronominal zero-ending (Ved. *ay-*, perhaps OIr. *é*), occurring together with a younger \*i-s (with nominal \*s- and zero-grade \*i- as in the f./n. and acc.). For the acc. m. cf. also Gr. Cypr. *iv*; the acc. f. is structurally reconstructed.

TABLE 1.33 IE DEMONSTRATIVE PRONOUN \*EY (\*IS) – \*IH<sub>2</sub> – \*ID (PLURAL)

pl.	PIE	Ved.	Lat.	OCS	Goth.
nom.	*e-y – *e-h <sub>2</sub> -es – *e-h <sub>2</sub>	im é –  im ás –  im á	ʌeī <sup>c</sup> – ʌeae\ – ʌea\	{oni – ony – ona} <sup>f</sup>	e s  –  ij os –  ij a
acc.	*e-o-ns <sup>a</sup> – *e-h <sub>2</sub> -(n) s – *e-h <sub>2</sub>	im ʌá n –  im ás –  im á	ʌeō s – ʌeā s – ʌea\	{je – je} – ja	ʌʌns –  ij os –  ij a
gen.	*e-y-som – *e-h <sub>2</sub> -som	ešʌám – āsʌám	{eum} <sup>d</sup> – ʌeārum\	ixъ	ʌʌz{e} – ʌʌzo
dat.	*e-y-b <sup>b</sup> /mos – *e-h <sub>2</sub> -b <sup>b</sup> /mos	ebh y ás – ābh y ás		imъ	ʌʌm
loc.	*e-y-su – *e-h <sub>2</sub> -su	ešú – āsú		ixъ	
instr.	*e-y-b <sup>b</sup> i(s) – *e-h <sub>2</sub> -b <sup>b</sup> i(s) <sup>b</sup>	ebhís – ābhís	[dat. {iis}] <sup>e</sup>	imi	

<sup>a</sup> Contracted to \*ōns; <sup>b</sup> BSL. \*-miHs; <sup>c</sup> Plautus (also ʌeī|s|), OLat. ʌeī|s|; <sup>d</sup> OLat., cf. Osc. *eisum|k|*; <sup>e</sup> OLat. {eieis}; <sup>f</sup> original forms in determinate adjective -ji, -je, -ja.

The \*ey- appearing in the plural is not the \*ey of the nom. sg. but \*e-y- (with pronominal \*-i), parallel to \*to-y- (see above), though synchronically it made no difference. The acc. pl. is difficult to reconstruct – Vedic seems to have a secondary *im-* (that later spread to the nom. pl. as well) by analogy to the acc. sg., the Latin forms are “de-contracted”, while Gothic has a secondary *ins* (with *i-* as in most forms).

### Interrogative pronouns

The interrogative pronoun for ‘who, what, which?’ has the root \*k<sup>w</sup>-, which shows up in two forms – as \*k<sup>w</sup>-o- and \*k<sup>w</sup>-i-. Cf. \*k<sup>w</sup>-o-s (m.) > Ved. *kás* ‘who, which?’ (GAv. sandhi-form *kas*), OCS *kъ-to* (oblique *ko-*), Lith. *kàs*, Goth. *hvas* (all ‘who?’); \*k<sup>w</sup>-o-d (n.) > Ved. *kád* ‘what, which?’ (GAv. *kať*), Lat. *quod* ‘which?’, OPruss. *ka* ‘what’, Goth. *hva* ‘what?’ (cf. \*k<sup>w</sup>-e-h<sub>2</sub> (f.) > Ved. *kā*, GAv. *kā*, Goth. *huo*). Cf. also \*k<sup>w</sup>-i-s (m./f.) > Ved. *ná-ṽk̄is* ‘no one’, GAv. *ciš*, Gr. *τίς*, Lat. *quis*, Hitt. *kuiš* (all ‘who?’); \*k<sup>w</sup>-i-d (n.) > Ved. *cid* ‘any’ (see below), GAv. *-cī* (pronominal particle), Gr. *τί*, Lat. *quid*, Slav. \*čь (Croat. Čak. *čä*, cf. OCS *čb-to*), Hitt. *kuit* (all ‘what?’). The relation of \*k<sup>w</sup>-i- and \*k<sup>w</sup>-o- is not clear; perhaps the first was originally ‘who/what?’, the second (adjectival) ‘which?’, cf. the opposition of Lat. *quis* ‘who?’ – *quid* ‘what?’ and *quod* ‘which?’, and Slav. \*čь ‘what’ and \*kъ ‘which’ (Croat. North Čak. *kä*, OCS *kyi* < \*kъ-jь). Alternatively, it can be supposed that the thematic \*k<sup>w</sup>-o- is innovative/secondary altogether (perhaps already in PIE). The declension is not easy to reconstruct, mainly because the relation of the two stems is unclear (\*k<sup>w</sup>-o- and \*k<sup>w</sup>-i-/k<sup>w</sup>ey-, cf., e.g., the acc. sg. Ved. *kám* < \*k<sup>w</sup>-o-m but Lat. *quem*, Hitt. *kuin* < \*k<sup>w</sup>-i-m ‘whom?’; nom. pl. Ved. *ké*, GAv. *kōi* < \*k<sup>w</sup>-o-y but GAv. *caīias*, OLat. *quēs* < \*k<sup>w</sup>ey-es), but the reconstructable forms apparently parallel those of the demonstrative \*so, cf. the gen. \*k<sup>w</sup>-e-so ‘of whom/what?’ (GAv. *cah|iī*ā, Gr. Ion. *τέο*, early Lat. *qu’o\|i|u|s|*, OCS *česo* ‘of what?’, Goth. *hvis*), dat. \*k<sup>w</sup>-o-sm-ōy ‘to whom?’ (Ved. *kásmai*, GAv. *kahmāi*, Umbr. *pusme*, OCS *komu*, Goth. *hvamma*), etc. It is unclear whether separate \*k<sup>w</sup>-o-so (Ved. *kás|y|a*, GAv. *kah|iī*ā ‘of whom/what?’, OCS *ko{g}o* ‘of whom?’), \*k<sup>w</sup>-e-sm-ōy (OCS *čemu* ‘to what?’), etc. existed (with the oblique suppletive \*k<sup>w</sup>e- in alternation with the nom./acc. \*k<sup>w</sup>i-, and the oblique \*k<sup>w</sup>-o- from the nom./acc. \*k<sup>w</sup>-o-), though they are formally easily reconstructable. For interrogative pronominal adverbs, cf. p. 89.

### Indefinite pronouns

The indefinite pronouns were derived from the interrogative ones. One pattern was doubling, cf. \*k<sup>w</sup>is-k<sup>w</sup>is ‘whoever’ (Lat. *quisquis*, Hitt. *kuiš kuiš*), \*k<sup>w</sup>os-k<sup>w</sup>id (Ved. *kás cid* ‘anyone’, GAv. *kascī* ‘every, any’), etc. The other one was adding the particle \*k<sup>w</sup>e (p. 105), cf. \*k<sup>w</sup>os-k<sup>w</sup>e (Ved. *kás ca* ‘anyone’, Arm. *ok<sup>c</sup>* ‘someone’ – p. 440), \*k<sup>w</sup>is-k<sup>w</sup>e (Lat. *quisque* ‘whoever’), \*k<sup>w</sup>od-k<sup>w</sup>e (Lat. *quodque* ‘whatever’, Goth. *hvah* ‘each, every’ (n.)), etc. Sometimes simple interrogative pronouns were probably used as indefinite pronouns as well.

### Relative pronouns

The relative pronoun is reconstructed as \*y-o (m.) – \*y-e-h<sub>2</sub> (f.) – \*y-o-d (n.) ‘who/what/which’ (sometimes thought to be a thematization of the dematerialized \*i-, p. 87), with the same declension as \*so – \*seh<sub>2</sub> – \*tod (p. 85–86). Cf. Ved. *yás* – *yá* – *yád*, Gr. *ὅς* – *ῆ* – *ὸ*, OCS *i|žel* (< \*jь) – *ja|žel* – *je|žel* (for *že* cf. p. 535), Clb. *io|s|* (acc. sg. *iom*, dat. sg.

*iomui*), Gaul. *dugiionti-io* ‘who serve’, OPhryg. *ios*. In many languages, the interrogative \*k<sup>w</sup>- has taken over the function of the relative \*yo- (e.g., in Anatolian, Tocharian, Italic), sometimes (as in part of Slavic) after the first attested texts. However, some doubt the antiquity/originality of the relative \*yo (cf. also p. 166).

### Pronominal adjectives

There are a number of pronominal adjectives, cf. \*h<sub>2</sub>elyos ‘other’ (Gr. ἄλλος, Lat. *alius*, Goth. *aljīs*, OIr. *aile*, Toch. B *alyek*), \*k<sup>w</sup>oteros ‘which of the two’ (OCS *kot|o|ryi*, Lith. dial. *kat|a|rās* ‘which’, and cf. p. 55, 80), etc.

### Pronominal adverbs

The pronominal adverbs are rather diverse in the IE languages, and not a lot of clear-cut correspondences exist. Cf. \*k<sup>w</sup>u-d<sup>h</sup>e ‘where?’ (Ved. *kūha*, OCS *kūde*; cf. Gr. πόθεν ‘whence?’ < \*k<sup>w</sup>o-d<sup>h</sup>e-n) and \*k<sup>w</sup>e/o-ti ‘how many?’ (Ved. *kāti*, YAv. *ca’ti*, Lat. *quot*, Bret. *pet*, cf. Gr. adj. πόσος < \*k<sup>w</sup>otyos) from \*k<sup>w</sup>- (p. 88); Ved. *i-tás* ‘from here’ from \*i- (p. 87), cf. Gr. ἐν-τός, Lat. *in-tus* ‘inside’ for the suffix \*-tos; similar forms from pronominal roots with the suffix \*-r: Lat. *quō-r* ‘why?’, Lith. *ku-ř* ‘where?’, Goth. *hva-r* ‘where?’, Alb. *ku-r* ‘when?’, Arm. *u-r* ‘where?’, etc.

### NUMERALS

PIE clearly had a decimal system. The reconstructed main cardinal numbers (with more or less direct reflexes), leaving out the many problematic issues, are:

1. \*h<sub>1</sub>oy-n-o-s ‘one’ (m.) (o-stem), \*-eh<sub>2</sub> (f.), \*-o-m (n.) > Gr. ὀνῆ ‘the ace on dice’, OLat. *oinos*, Lith. *viėnas* (cf. Slav. \*īnъ ‘other’ < \*h<sub>1</sub>eyn-, OCS *in-o-rogb* ‘unicorn’), OPruss. acc. sg. *ainan*, Goth. *ains*, OIr. *óen* (with different suffixes Ved. *ékas*; OPers. *aiva* ‘one’, Gr. Cyp. οἴφος ‘only’, p. 49); \*sōm (m.) ‘one’ (OCS *sam-ъ* ‘alone’) – \*sm-ih<sub>2</sub> (f.) (Gr. μία, Arm. *mi*, cf. p. 86 for \*si-) – \*sem (n.) (Gr. ἕν, cf. Lat. *sem-el* ‘once’), cf. Ved. *sa-kṛt* ‘once’ (\*sm̥-), Gr. ὁμός ‘one and the same’, Goth. *sama* ‘same’ (\*som-)
2. \*d(u)w-o-h<sub>1</sub> (m.) ‘two’ (cf. p. 54 for Lindemann’s Law; o-stem du.) > Ved. *d(u)vā* (-*āu*), Gr. Hom. δύο, Lat. *duo*, OCS *dъva*, Lith. *dù*, OIr. *dóu*, Arm. *erku* (p. 433–434), Toch. A *wu*; \*d(u)w-eh<sub>2</sub>-ih<sub>1</sub> (f.) (eh<sub>2</sub>-stem du.), \*d(u)w-o-yh<sub>1</sub> (n.) (o-stem du.) > Ved. *d(u)vé* (f./n.), Lat. *duae* (f.), OCS *dъvě* (f./n.), Lith. *dvi* (f.), Toch. A *we* (f.); cf. likewise \*-b<sup>h</sup>-o-h<sub>1</sub> ‘both’ (Goth. *b{ai}*), \*h<sub>2</sub>ent-b<sup>h</sup>oh<sub>1</sub> (Toch. B *antapi*, Gr. ἄμφω, Lat. *ambō*), also OCS *oba*, Lith. *abù* (\*h<sub>2</sub>e-b<sup>h</sup>oh<sub>1</sub>?), Ved. *ubhā* (\*Hu-b<sup>h</sup>oh<sub>1</sub>?)
3. \*tr-ey-es ‘three’ (m.) (i-stem, p. 32) > Ved. *tráyas*, Gr. τρεῖς, Lat. *trēs*, OCS *trije*, Lith. *trỹs*, Goth. *þreis*, OIr. *trī*, Arm. *erek*, Toch. B *tr’āvi*, cf. loc. \*tr-i-su (Ved. *trishū*), etc.; \*tisres (f.) (< \*tri-sr-es, r-stem) > Ved. *tisrás*, OIr. *teoir*; \*tr-ih<sub>2</sub> (n.) (i-stem pl.) > Ved. *trī*, Gr. τρία, OCS *tri*
4. \*k<sup>w</sup>etwor-es ‘four’ (m.) (r-stem, p. 23) > Ved. *catvāras*, Gr. τέτταραρες, OCS *četvyre*, OIr. *ceth(a)ir*, Arm. *čork*, Toch. B *štwer*; oblique/compound \*k<sup>w</sup>tw̥rC-/urV- (p. 52) > Arm. *k’ar-*; \*k<sup>w</sup>ete-sr-es (f.) (< \*k<sup>w</sup>etwer-sr-es?) > Ved. *cātasras*, OIr. *cetheoir*; \*k<sup>w</sup>etwōr (n.) < \*-orh<sub>2</sub> (p. 70) > Lat. *quārtuor*, Goth. *ƿīdwor* (reshaped \*k<sup>w</sup>etwor-h<sub>2</sub> (n.) > Ved. *catvāri*, Gr. τέτταρα)

5. \*penk<sup>w</sup>e ‘five’ (p. 24) > Ved. *pāñca*, Gr. πέντε, Lat. *quīnque*, Lith. *penk{i}*, Goth. *fiṃf*, OIr. *cóic*, Arm. *hing*, Alb. *pesë*, Toch. A *pāñ*, B *piś*; cf. \*pṛk<sup>w</sup>-stis ‘fist’ > PSI. \*pṛṣṭṣ (BCMS *pēst*), OEng. *fýst*
6. \*sw/s/weks ‘six’ > Ved. *śát*, Gr. Dor. ἑξήκοντα, Lat. *sex*, Goth. *saihs*, OIr. *sé* (W *chwech*, Gaul. *suexos*), Arm. *vec* (p. 440), Toch. B *śkas*
7. \*septm ‘seven’ (loanword from Semitic \*šabṣ-at-um, p. 29) > Ved. *saptá*, Gr. ἑπτὰ, Lat. *septem*, Lith. *sept{yni}*, Goth. *sīḅuṽn*, OIr. *secht<sup>n</sup>*, Arm. *ewt<sup>n</sup>*, Toch. A *špät*
8. \*h<sub>3</sub>ek<sup>tō</sup>(w) ‘eight’ (p. 45) > Ved. *aṣṭā* (-áú), Gr. ὀκτώ, Lat. *octō*, Lith. *aštuo{n}*, Goth. *ahtau*, OIr. *ocht<sup>n</sup>*, Arm. *uṭ<sup>c</sup>* (p. 440), Toch. B *okt*
9. \*h<sub>1</sub>new<sup>ṇ</sup> ‘nine’ (p. 35) > Ved. *náva*, Gr. ἐννέα, Lat. *nouem*, Goth. *niun*, OIr. *noi<sup>n</sup>*, Arm. *inn*, Toch. AB *ñu*
10. \*dek<sup>m</sup> ‘ten’ (p. 34) > Ved. *dáśa*, Gr. δέκα, Lat. *decem*, Goth. *taihun*, OIr. *deich<sup>n</sup>*, Arm. *t<sup>h</sup>áśn*, Toch. B *śak*
20. \*wīk<sup>n</sup>tih<sub>1</sub> < \*dw-i(h<sub>1</sub>)-dk<sup>m</sup>t-ih<sub>1</sub> ‘twenty’ (du.) > Ved. *viṃśatīś*, YAv. *visa<sup>t</sup>i*, Gr. Dor. *ἑκατὼν*, Lat. *uīginti*, OW *uceint*, Arm. *k<sup>c</sup>san*, Toch. A *wiki*
100. \*k<sup>m</sup>tom < \*dk<sup>m</sup>t-o-m ‘hundred’ (p. 34) > Ved. *śátám*, Gr. ἑκατόν, Lat. *centum*, OCS *st<sup>o</sup>*, Lith. *šimtaś*, Goth. *hunda*, Clb. *kantom*, Toch. B *kante*
1000. \*g<sup>hesl</sup>- ‘thousand’ > Ved. *śaśān*, Gr. Aeol. χελλιοι (Lat. *mīlle* < \*smih<sub>2</sub>-g<sup>hesl</sup>- ‘one thousand’); dial. \*tuHs-(d)k<sup>m</sup>t-ih<sub>2</sub> (literary \*‘strong hundred’) > OCS *tyse<sup>o</sup>št<sup>i</sup>*, OPruss. acc. pl. *tūsint{ons}*, Goth. *þusundi* (Eng. *thousand*)

Only numbers 1–4 had declension and distinguished gender. The two forms for ‘one’ (disregarding the different suffixes with the root \*h<sub>1</sub>oy-) perhaps had slightly different semantics or usage. In 3 and 4, Indo-Iranian and Celtic point to feminine forms with \*-sr-, often interpreted as ‘woman’ (cf. \*swe-sr- ‘sister’, etc.). The word for ‘twenty’ is originally ‘two tens’, and ‘hundred’ is derived from ‘ten’ (‘ten’ itself is usually thought to be ‘two-hand’, cf. Goth. *handus* ‘hand’ < \*k<sup>o</sup>mt-). The numbers from 11 to 19 were probably made by the pattern ‘one-ten’, ‘two-ten’, etc., cf. \*dwoh<sub>1</sub>-dek<sup>m</sup>(t) ‘twelve’ > Ved. *divádaśa*, Gr. δώδεκα, Lat. *duodecim*. The decades 30–90 were made with \*-(d)k<sup>o</sup>mt<sub>2</sub> (originally \*-(d)k<sup>o</sup>mt with the loss of \*-h<sub>2</sub>?), the old collective of \*dek<sup>m</sup>(t), cf. \*tri-h<sub>2</sub>-(d)k<sup>o</sup>mt<sub>2</sub> ‘thirty’ > Gr. τριᾶκοντα, Lat. *trīgintā*, OIr. *tricho*, Arm. *eresun* (p. 440–441), Toch. B *lāryak<sup>h</sup>* (Toch. B *-ka* could be from \*-(d)k<sup>o</sup>mt as well, cf. p. 467), etc. The numbers 21–29, 31–39, etc. were perhaps made the same way as 11–19, cf. \*treys-trih<sub>2</sub>-(d)k<sup>o</sup>mt<sub>2</sub> ‘thirty-three’ (lit. ‘three-thirty’) > Ved. *trāyas-triṃśat*, etc. The hundreds from 200 to 900 are difficult to reconstruct but perhaps were (at least in one variant) compounds with \*-(d)k<sup>o</sup>mt- (cf. Gr. dial. -κατ-οι, Lat. *-cent-ī*).

The ordinal numbers were derived from the cardinal ones (except ‘one’) with various thematic suffixes. The original pattern is sometimes difficult to discern. The provisory reconstructions are:

1. \*pṛh<sub>2</sub>-w/m/t-o-s ‘first’ (< ‘foremost’) > Ved. *pūrvas* ‘fore, first’, OCS *pr<sup>o</sup>v<sup>o</sup>*, Toch. B *parwe(-šše)*; Lith. *pirmas*, Goth. *f<sup>r</sup>urum<sup>a</sup>*; Gr. Dor. *πρῶτος* (cf. adverbs \*p<sup>r</sup>eh<sub>2</sub>-i > Lat. *prae* ‘in front of’, \*p<sup>r</sup>h<sub>2</sub> > Gr. *παρά/πάρα* ‘beside’)
2. ‘second’ cannot be reconstructed (all languages have different roots)
3. \*t<sup>r</sup>-tri-t-y-o-s ‘third’ > Ved. *tṛtī<sup>y</sup>as*, YAv. *θritiia-*, Lat. *tertius*, OCS *tr<sup>e</sup>l<sup>e</sup>tii*, Lith. *tr<sup>e</sup>l<sup>e</sup>čias*; Gr. *τρίτος*, Goth. *þrid<sup>a</sup>*
4. \*k<sup>w</sup>tw<sup>y</sup>-t-o-s ‘fourth’ > Ved. *turī<sup>y</sup>as*, YAv. *tu<sup>r</sup>iia-*, Gr. *τετάρτος*, Lat. *quārtus*, OCS *č<sup>e</sup>l<sup>e</sup>tv<sup>o</sup>rb<sup>o</sup>*, Lith. *k<sup>e</sup>l<sup>e</sup>tvirtas*, OEng. *f<sup>e</sup>l<sup>e</sup>orþ<sup>a</sup>*, Toch. B *štarte*
5. \*penk<sup>w</sup>-t-o-s ‘fifth’ > Gr. *πέμπτος*, older Lat. *quinctus*, OCS *p<sup>e</sup>rb<sup>o</sup>*, Lith. *peñktas*, Goth. *fiṃfta-*, Toch. B *piñkte*

6. \*(s)(w)ek(s)-t-o-s ‘sixth’ > Ved. *ṣaṣṭ|h|ás*, Gr. ἕκτος, Lat. *sextus*, OCS *š\estь*, Lith. *š\ėštas*, OPruss. *uschts* (< \*ukstos), Goth. *saihts\á\*, Toch. B *škaste*
7. \*septm-o-s ‘seventh’ > Ved. *sap\am\ás*, Gr. ἑβδόμοος, Lat. *sept\i\mus*, OCS *sedmь*, OPruss. *septmas*; \*septm-t-o-s > Ved. *saptát|h|as*, Lith. *septiñtas*, OHG *sibunt\o\*, Toch. A *šäptänt*
8. \*h<sub>3</sub>ekt-ow-o-s > Gr. ὀκτώοος, Lat. *oct\ā\uus*
9. \*h<sub>1</sub>newn-t-o-s ‘ninth’ > Gr. ἔνατος, OCS *d\evetь*, Lith. *deviñtas*, Goth. *niund\á\*, Toch. B *ñunte*
10. \*dekmt-o-s ‘tenth’ > Gr. δέκατος, OCS *desetь*, Lith. *dešimtas*, Goth. *taihund\á\*, Toch. B *škante*; \*dekmt-H-o-s > Ved. *daśamás*, Lat. *decimus*

Multiplicative adverbial numerals \*dwis ‘twice’ (Ved. *dvís*, Gr. δίς, Lat. *bis*, Oic. *tvís-var*), \*tris ‘thrice’ (Ved. *trís*, Gr. τρίς, Lat. *ter*, Oic. *pris-var*), \*k<sup>w</sup>eturs ‘four times’ (Ved. *cátur*, Lat. *qu\á\ter*) can also be reconstructed (cf. also Arm. – p. 441).

## VERBS

The reconstruction of the PIE verbal system is more complex than the reconstruction of the nominal system. There is some concord when it comes to the reconstruction of specific forms and paradigms (which is the primary task here – determining the exact function of specific forms is more difficult), although it is not always clear what categories were present in what phase of PIE. There are a number of reasons for disagreement in the reconstruction of the PIE verbal system. First of all, the verbal systems of the early attested IE languages are much less congruent than the earliest nominal systems – e.g., while Vedic has three voices (active, middle, passive), five tenses (present, aorist, imperfect, perfect, future), and five moods (indicative, injunctive, imperative, subjunctive, optative), Hittite has only two of each (active, middle; present, preterite; indicative, imperative). Thus, the PIE system is more difficult to reconstruct, and there is more room for different possible reconstructions. Generally speaking, analyzing verbal systems is often hard and controversial even when it comes to living languages, so it is no wonder that this is even more difficult when it comes to dead languages (like Vedic) and even more so for reconstructed languages (like PIE).

The earliest, traditional reconstructions of the PIE verbal system were based on the so-called Greco-Aryan model and the rich verbal morphologies of (mainly) Greek and Old Indic. However, as became apparent after the discovery of Anatolian (and Tocharian, but to a smaller extent), this reconstructed system was very different from the less complex systems found in Anatolian languages – for instance, Anatolian shows no traces of the subjunctive and optative moods that have to be reconstructed for the rest of IE. The basic problem is how to interpret these discrepancies – to consider them as the result of post-Anatolian common innovations in late PIE, to consider Anatolian as innovative (e.g., losing the subjunctive and optative), or some kind of combination of both (p. 194). This then results in the fluid nature of what exactly is reconstructed for PIE – sometimes it is Early PIE (with Anatolian) and sometimes it is Late PIE (post-Anatolian).

The following PIE verbal categories can be reconstructed (we will mention possible others later in the text):

NUMBER: singular, dual, plural  
PERSON: 1st, 2nd, 3rd



ASPECT: present, aorist, perfect  
 VOICE: active, middle  
 MOOD: indicative, imperative, optative, subjunctive  
 NOMINAL VERB FORMS: participles (and verbal adjectives, verbal nouns)

Things are mostly straightforward when it comes to number, person, voice, and nominal formations. What were traditionally regarded as tenses are now most often interpreted as aspects (the present being imperfective, the aorist perfective, and the perfect stative aspect),<sup>8</sup> with the system best preserved in (early) Greek. Thus, for PIE we can reconstruct the imperfective (“present”) \*stisteh<sub>2</sub>mi ‘I am getting up’ – the perfective (“aorist”) \*steh<sub>2</sub>m ‘I stood up’ – the stative (“perfect”) \*stestoh<sub>2</sub>e ‘I am standing’ (not all verbs could form all three). Present/imperfective and aorist/perfective are grouped together as eventive (as opposed to perfect, i.e., stative). The categories of voice and mood existed in the eventive system only (thus, the perfect had no middle and no moods). Both the present and aorist had separate middle forms. The (active or middle) indicative is unmarked, and the rest of the moods functioned as separate simple formations (there was only one optative, subjunctive, and imperative – though in the case of the latter also with special future endings); i.e., there were no separate forms for the subjunctive present, subjunctive aorist, etc. as in Vedic or Greek. The present/aorist opposition, optative, subjunctive, and dual (and special formations like thematic presents or reduplicated perfects) did not exist in Anatolian and were possibly post-Anatolian.

All verbal roots had a CVC structure (with possible clusters – p. 52–53), with \*e usually thought to be the “basic” ablaut grade (p. 43), cf. \*b<sup>h</sup>er- (p. 40), \*d<sup>h</sup>eh<sub>1</sub>- (p. 45), though there were exceptions like \*kan- (p. 42). PIE verbal forms had the following structure: root (+ affix) (+ thematic vowel) + ending (+ particle). A verb had to have a root and an ending (even if a zero morpheme). Some forms also had derivational suffixes (a thematic vowel was a particular kind of them) and end-particles. Cf. verbal forms like \*h<sub>1</sub>s-te ‘you (pl.) are’ (root + ending), \*b<sup>h</sup>er-e-te ‘you (pl.) are carrying’ (root + thematic vowel + ending), \*b<sup>h</sup>owd<sup>h</sup>-ey-e-me-s ‘we are waking (somebody)’ (root + suffix + thematic vowel + ending + particle). Reduplication (cf. \*d<sup>h</sup>e-d<sup>h</sup>eh<sub>1</sub>-mi ‘I am putting’) was more frequent than in nouns, and there was also an infix (cf. \*yu-ne-g-mi ‘I am yoking’ from the root \*yug-/yewg-).

PIE had a special class of particles that were attached after endings in some finite verbal forms. Some were a sort of secondary markers for a certain category, others had no clear function, some were probably facultative, and some forms could have had more than one particle. The major verbal particles were \*-i (the present, *hic & nunc*, particle), \*-u (the imperative particle), and \*-o and \*-r (the middle particles). Minor ones were the 1 (and 2) pl. and 1/2/3 du. present \*-s and \*-n, and the 1 pl. middle \*-d<sup>h</sup><sub>2</sub> (others also occur in specific languages or cases).

## Present

What is traditionally called the “present” in world languages often does not only stand for events in real present time (which is its primary and prototypical role) but can also be used for general statements, future, past, etc. This was undoubtedly the case in PIE as well – for instance, having no separate future tense, PIE surely had to “replace” it with present forms (at least in some instances). However, we cannot get into the specifics of the usage of the present in PIE here.

There were two basic present conjugations (similar to noun declensions): thematic (with the thematic vowel \*-e/o- between stems and endings and with no ablaut alternation) and

athematic (with endings right after the stems, and with ablaut alternations – most often \*e in sg. and \*Ø in pl.), like \*b<sup>h</sup>er-e-te ‘you are bearing (pl.)’ but \*h<sub>1</sub>d-te ‘you are eating (pl.)’. Thematic and athematic verbs had the same/similar endings in the plural (\*-me-, \*-te, \*-nt-), while the singular forms originally differed. Athematic (and partly thematic) endings \*-m-, \*-s-, \*-t- (sg.), \*-me-, \*-te, \*-nt- (pl.) are shared with the aorist (p. 95–97). Some endings (especially athematic ones) had a *hic & nunc* particle \*-i, while the 1 pl. had also \*-s or \*-n (also sometimes in the 2 pl.) – the endings with these present particles are traditionally called “primary” (though they are not really primary). In later IE languages, the non-ablauting thematic forms usually win, but usually with incorporation of at least some originally athematic endings. Here is the PIE athematic paradigm with reflexes:

TABLE 1.34 IE ATHEMATIC PRESENT

	PIE	Ved.	Gr.	Lat.	OCS	OLith.	Goth.	Hitt.
1 sg.	*h <sub>1</sub> es-m-i	<i>ásmi</i>	εἰμί	{ <i>sum</i> }	<i>jesmъ</i>	<i>esm</i> {i} <sup>e</sup>	<i>im</i>	<i>ēšmi</i>
2 sg.	*h <sub>1</sub> e(s)-s-i	<i>ási</i>	εἶ	<i>es</i>	<i>jes</i> {i} <sup>e</sup>	<i>es</i> {i} <sup>e</sup>	<i>is</i>	<i>ēšši</i>
3 sg.	*h <sub>1</sub> es-t-i	<i>ásti</i>	ἔστί	<i>est</i>	<i>jest</i> ʌ\	<i>ēsti</i>	<i>ist</i>	<i>ēšzi</i>
1 pl.	*h <sub>1</sub> s- me/o(-s/n-(i))	<i>smási</i> <sup>b</sup>	Hom. εἰμέν	{ <i>sumus</i> }	{ <i>jesmъ</i> }	{ <i>esm</i> ʌ\}	{ <i>sijum</i> }	{ <i>ešuwʌni</i> } <sup>g</sup>
2 pl.	*h <sub>1</sub> s-te	<i>st/h</i> a <sup>c</sup>	ἔστέ	{ <i>esti</i>  s }	{ <i>jest</i> e}	{ <i>est</i> ʌ\}	{ <i>sijub</i> }	<sup>h</sup>
3 pl.	*h <sub>1</sub> s-ent-i <sup>a</sup>	<i>sánti</i>	Dor. ἐντί	sʌʌnt <sup>d</sup>	sʌʌʌnt <sup>f</sup>	{ <i>ēsti</i> }	<i>sind</i>	<i>ašʌʌnzi</i>

<sup>a</sup> With a younger variant \*-onti; <sup>b</sup> also *smas*; <sup>c</sup> also *sthá-na*; <sup>d</sup> Umbr. *sent*; <sup>e</sup> with middle \*-oy (p. 99) instead of \*-i; <sup>f</sup> Slav. also \*sęts ( > Mac. *ce*), OCS *dadęts*; <sup>g</sup> accented *-weni*; <sup>h</sup> *-teni*.

Athematic presents had a full grade (\*h<sub>1</sub>es-) in the sg. and a zero grade (\*h<sub>1</sub>s-) in the pl. Most endings had the present particle \*-i. In the 1 pl. (like in thematic stems), numerous variants seem to have existed (the ending had ablaut variants \*-me/o, together with particles \*-s or \*-n, and the present particle \*-i in some languages): \*-mes (Gr. Dor. -μες, Ved. *-mas*, Cz. *-me*), \*-mesi (Ved. *-masi*), \*-men (Gr. Att. -μεν), \*-mo (Slav. *-mo*), \*-mos (Lat. *-mus*, OCS *-mъ*, Goth. *-m*), \*-meni (Hitt. ʌʌʌeni), etc. (in some cases, different developments are possible). Their original distribution, exact formation, and number of existing variants is unclear. The athematic 3 pl. \*-enti was replaced early on with thematic \*-onti in many languages (perhaps facultatively/dialectally already in PIE). The thematic endings were:

TABLE 1.35 IE THEMATIC PRESENT

	PIE	Ved.	Gr.	Lat.	OCS	Lith.	Goth.
1 sg.	*-o-h <sub>2</sub>	-ā mi	-ω	-ō	-ʌʌ	-ù	-a
2 sg.	*-e-h <sub>1</sub> -i		-ε ç			-i <sup>e</sup>	
	*-e-s-i	<i>-asi</i>		<i>-is</i>	-eʌʌ{i} <sup>c</sup>		<i>-is</i>
3 sg.	*-e		-ε t		<sup>d</sup>	-ʌʌ	
	*-e-t-i	<i>-ati</i>		<i>-it</i>	-eʌʌ\		<i>-iþ</i>
1 pl.	*-o-me/o	<i>-āmas(i)</i>	-ομεν <sup>b</sup>	<i>-imus</i>	-ʌʌmo	-ʌʌmʌe\ <sup>f</sup>	<i>-am</i>
	(-s/n-(i))						
2 pl.	*-e-te <sup>a</sup>	<i>-at/h</i> a	-ετε	<i>-iti</i>  s	<i>-ete</i>	-ʌʌʌe\ <sup>f</sup>	<i>-iþ</i>
3 pl.	*-o					-a	
	*-o-nt-i	<i>-anti</i>	Dor. -οντι	<i>-unt</i>	-ʌʌʌ\		<i>-and</i>

<sup>a</sup> Cf. Toch. A *-äc*; <sup>b</sup> Dor. -ομες; <sup>c</sup> Slav. also \*-ešb; <sup>d</sup> -e in West and (modern) South Slavic; <sup>e</sup> reflexive *-ie-si*; <sup>f</sup> reflexive *-mē-si*, *-tē-si* < \*-mē, \*-tē (cf. perhaps the rare Ved. scansion *-mā* in the aor. 1 pl.).

The 1 sg., 1 pl., and 3 pl. had thematic \*-o-, and the other forms had \*-e-. In the 1 sg., the reflexes point to \*-ō, which is interpreted as \*-oH (on the assumption that there were no \*-V̄# type endings<sup>9</sup>) and further as \*-o-h<sub>2</sub> (on the basis of the supposed relation to the perfect ending \*-h<sub>2</sub>e and middle \*-h<sub>2</sub>(e) – p. 98–99), which is frequently reconstructed, though in fact speculative. The usual reconstructions for the 2/3 sg. and 3 pl. are \*-esi, \*-eti, \*-onti (attested in most languages), with the endings as in athematic stems. In contrast to what is usually reconstructed, one can suppose that thematic stems originally had different endings, perhaps coexisting with younger athematic-influenced ones already in the last stage of PIE. The heterodox thematic reconstruction \*-eh<sub>1</sub>i in the 2 sg. is based on Greek (if not from \*-esi + \*-s) and Lithuanian (the accent points to a laryngeal); \*-e in the 3 sg. again on Greek/Balto-Slavic (with analogical thematic \*-o in Lithuanian) and Toch. B *-ā-m* (< \*-e-nu),<sup>10</sup> \*-o in the 3 pl. on Lithuanian (where sg. = pl.) and Toch. A *-e* < \*-o (together with *-eñc* < \*-onti) – cf. Beekes 2011: 260. The original thematic endings in the 3 sg./pl. would thus be bare thematic vowels in different ablaut. Anatolian had no thematic present.

The endings of the dual (athematic/thematic) are not easily reconstructable:

TABLE 1.36 IE PRESENT DUAL ENDINGS

	PIE	Ved.	Gr.	OCS	Lith.	Goth.
1 du.	*-(o)-we-(s)	-(ā)vas <sup>b</sup>		-(e)νλ̥	-(a)νλ̥a	-(o)s
2 du.	*-(e)-tHV-(s/n)	-(a)thas	-(ε)των	-(e)ta	-(a)ta <sup>c</sup>	-ts
3 du.	*-(e)-te/o-s/n <sup>a</sup>	-(a)tas	-(ε)των	-(e)te		

<sup>a</sup> Cf. Toch. B *-tem* < \*-ton-V; <sup>b</sup> cf. Av. *-uuahi*; <sup>c</sup> cf. perhaps Umbr. 2 pl. imp. *eta-to* < \*-tā.

Cf. the pronoun \*wē ‘us two’ (p. 84) for 1 du. Ved. 2 pl. *-tha* is perhaps influenced by the 2 du.

The formation of the present stem is much more diverse than in other verbal forms – both thematic and athematic present conjugations were formed in numerous ways (usually with different affixes – some with a specific derivational meaning, others without it or with a meaning difficult to reconstruct). The list here is non-exhaustive and simplified:

THEMATIC – (a) root-stems (cf. PIE \*weǵ<sup>h</sup>-e-, \*b<sup>h</sup>er-e- above; with zero-grade PIE \*g<sup>w</sup>ṛh<sub>3</sub>-e- ‘devour’ > Ved. *girāti*, OCS *žbrǫ*); (b) reduplicated stems (PIE \*pi-ph<sub>3</sub>-e- from \*peh<sub>3</sub>- ‘drink’ > Ved. *pibāmi*, Lat. *bibō*); (c) \*-y- suffix (PIE \*leh<sub>2</sub>-y-e- ‘bark’ > Ved. *rāyati*, OCS *lajǫ*, Lith. *lōju*; also in denominatives like Ved. *apasyāti* ‘works’ from *apas-* ‘work’ and rare causatives like Lat. *sōpiō* ‘I put to sleep’ < \*swōp-y-e from \*swep- ‘sleep’); (d) \*-ey- causatives/iteratives<sup>11</sup> (PIE \*mon-ey-e- ‘make somebody think’ from \*men- ‘think’ > Lat. *moneō* ‘I remind’, Ved. *mānāyati* ‘honors’); (e) \*-sk- (PIE \*p<sup>r</sup>k-sk-e- ‘ask’ > Ved. *p<sup>r</sup>cc<sup>h</sup>āti*, Lat. *poscō*); (f) \*-eh<sub>1</sub>- staves (PIE \*sed-eh<sub>1</sub>-e- ‘to be seated’ from \*sed- ‘sit’ > Lat. *sedēre*, OCS *sēdēti*, Lith. *sēdėti* ‘to sit’); (g) \*-eh<sub>2</sub>- factitives (PIE \*new-eh<sub>2</sub>-e- ‘to make new’ > Hitt. *nēwah<sup>h</sup>*- from \*newos ‘new’); (h) \*(e)s- (Lat. *gerō* ‘I carry’ < \*h<sub>2</sub>ǵ-es-e- from \*h<sub>2</sub>eǵ- > Lat. *agō* ‘I drive’; also desiderative \*-s-, which became the new future in many IE languages)

ATHEMATIC – (a) root-stems (cf. PIE \*h<sub>1</sub>es- above or \*h<sub>1</sub>ey- ‘go’ > Ved. *émi*, Gr. εἶμι; “Narten-presents” with long root, cf. Ved. *dāṣṭi* ‘waits’ < \*dēk-ti); (b) reduplicated stems (PIE \*de/i-deh<sub>3</sub>- ‘give’ > Ved. *dādāmi*, Gr. δίδωμι); (c) \*-n(e)- infix (PIE \*yu-n(e)-g- ‘yoke’ > Ved. *yunákti*, pl. *yuñjánti*, Lat. *iungō*); (d) \*-new-/nu- suffix (Ved. *kr̥nómi* ‘I make’, *kr̥nuthá* ‘you (pl.) make’)

One can distinguish primary (derived directly from a verbal root without a suffix or with a suffix with no exact derivational meaning) and secondary formations (derived with a suffix that changed the meaning of the verbal root – e.g., causative, iterative, etc. – or from a noun or adjective root). The primary ones are all athematic and thematic (a)-(c). One verb could have had more than one coexisting primary present forms (with or without different connotations) – cf. from \*g<sup>w</sup>em- ‘go, come’ both \*g<sup>w</sup>ṃ-sk-e- (Ved. *gáčhati* ‘goes, comes’, Gr. βάσκε! ‘come!’) and \*g<sup>w</sup>ṃ-y-e- (Gr. βαίνω ‘I walk’, Lat. *ueniō* ‘I come’). It is possible, even likely, that some assumed thematic formations were originally athematic. What the exact formation of a specific PIE verb was has to be reconstructed word by word (cf. LIV). For the middle present see p. 99–100.

### Aorist

The PIE aorist denoted perfective aspect. Its endings were basically the same as those in the athematic present (p. 93), pointing probably to a common pre-PIE origin of the present and aorist. However, the aorist did not have particles attached to endings as in the present (particularly not the *hic & nunc* \*-i), except in the dual (but different from the present dual). The endings in the aorist (traditionally called “secondary”) are thus actually unmarked (unlike the present ones). Vedic preserves the complete opposition of the present and aorist endings (cf. Ved. pres. *-mas(i)* but aor. *-ma*), while many languages later generalized the original present endings in the aorist as well (cf. Gr. -μεν in pres./aor.). Greek and Indo-Iranian are the only branches to preserve both the aorist and perfect. In Italic and Celtic the aorist merges with the perfect into new preterite formations; in Germanic it is lost; in Slavic, Armenian, and Albanian it is preserved (while the perfect is generally lost there except for traces in Balto-Slavic and Armenian). Anatolian has no aorist (which is perhaps not an innovation). Greek, Indo-Iranian, Phrygian, and (limitedly) Armenian have a so-called augment in the past tenses (including the aorist), attached to the beginning of the form (Ved. *a-*, Gr. ἐ-, Phryg./Arm. *e-*), which stems from the PIE adverb \*h<sub>1</sub>e ‘at that time’. Some linguists reconstruct the augment as part of the aorist for PIE as well (where it may have been at the most a facultative or dialectal feature).

Unlike the numerous present formations, the aorist had only these basic formations:

- 1) root-aorist – root [sg. – full grade, pl. – zero grade] + endings \*deh<sub>3</sub>-m – pl. \*dh<sub>3</sub>-me
- 2) sigmatic aorist – root [long grade] + \*-s- + endings \*wēǵ<sup>h</sup>-s-ṃ
- 3) thematic aorist – root [zero grade] + thematic vowel + endings \*wid-o-m

The root-aorist is attested in Indo-Iranian, Greek, and Armenian (perhaps partially in Old Church Slavic) only but is often considered the most archaic (while sigmatic/thematic aorists are often considered secondary). It is formed from the endings attached directly to the root (often of those verbs that have a reduplicated athematic present), with the \*e

grade in the sg. and the zero grade in the pl. (the original ablaut is preserved in some verbs in Greek):

TABLE 1.37 IE ROOT-AORIST

	PIE	Ved.	Gr.
1 sg.	*deh <sub>3</sub> -m <sup>a</sup>	<i>dām</i> <sup>b</sup>	ἔ δω {κα} <sup>d</sup>
2 sg.	*deh <sub>3</sub> -s	á  <i>dās</i>	ἔ δω {κας}
3 sg.	*deh <sub>3</sub> -t	á  <i>dāt</i>	ἔ δω {κε}
1 pl.	*dh <sub>3</sub> -me	a  <i>dāṃma</i>	ἔ δομε v
2 pl.	*dh <sub>3</sub> -te	<i>sth'āta</i>	ἔ δοτε
3 pl.	*dh <sub>3</sub> -nt	á  <i>d</i> {ur} <sup>c</sup>	ἔ δο {σαν} <sup>e</sup>

<sup>a</sup> Asyllabic because of Stang's Law (cf. p. 70); <sup>b</sup> cf. *ápām* 'has drunk'; <sup>c</sup> cf. the original *a-bhūv-an* < \*b<sup>h</sup>uh<sub>2</sub>-nt, Av. -at < \*-nt; <sup>d</sup> cf. the original ἔ-στη-v 'stood' < \*steh<sub>2</sub>-m; <sup>e</sup> cf. the more original (Homeric) ἔστ-αν < \*sth<sub>2</sub>-nt

Cf. also \*steh<sub>2</sub>m 'stood' (Ved. *ásthām*, Gr. ἔστην), \*d<sup>h</sup>eh<sub>1</sub>m 'put' (Ved. *adhām*, Gr. ἔθηκα with secondary -κα, 1 pl. ἔθεμεν, Arm. 3 sg. *ed*, etc. The root was perhaps exceptionally zero grade in \*b<sup>h</sup>uh<sub>2</sub>- 'be': 3 sg. \*b<sup>h</sup>uh<sub>2</sub>t (Ved. *ábhūt*, Gr. ἔφῶ, (?) OCS *by*). In Vedic and Homeric Greek, one can also find augmentless aorists like Ved. *bhūt*, Hom. φῶ 'became', etc.

The sigmatic aorist, very productive in late PIE and later languages, was formed from a long-grade root, the suffix \*-s-, and endings:

TABLE 1.38 IE SIGMATIC AORIST

	PIE	Ved.	Gr.	OCS
1 sg.	*wēǵ <sup>h</sup> -s-m̄	á  <i>prākṣa m</i>	ē ζευξα	<i>věs b </i>
2 sg.	*wēǵ <sup>h</sup> -s-s <sup>a</sup>	a  <i>yā's</i> <sup>b</sup>	ē ζευξ {ας}	
3 sg.	*wēǵ <sup>h</sup> -s-t	a  <i>prāṭ</i>	ē ζευξ {ε}	<sup>f</sup>
1 pl.	*wēǵ <sup>h</sup> -s-me	á  <i>jaiṣma</i> <sup>c</sup>	ē ζευξ α με v	<i>věs o m b </i>
2 pl.	*wēǵ <sup>h</sup> -s-te	á  <i>cchānta</i> <sup>d</sup>	ē ζευξ α τε	<i>věste</i>
3 pl.	*wēǵ <sup>h</sup> -s-nt	á  <i>vāks</i> {ur} <sup>e</sup>	ē ζευξ α v	<i>věse</i>

<sup>a</sup> Simplified to \*wēǵ<sup>h</sup>s probably already in PIE (p. 53); <sup>b</sup> 'has sacrificed'; from *yaj-* < \*Hyaǵ-; <sup>c</sup> 'have conquered'; from *ji-*; <sup>d</sup> < \*áčchantsta 'have seemed', from *chand-*; <sup>e</sup> cf. Av. -at < \*-nt; <sup>f</sup> cf. perhaps OCS 2/3 *by-st-b* 'were/was'.

The long grade in the sg./pl. is well attested in Vedic, Old Church Slavic (*věsṭ* 'conveyed' from *vez-* < \*weǵ<sup>h</sup>-, *basṭ* 'jabbed' from *bod-* < \*b<sup>h</sup>od<sup>h</sup>-, *grěsṭ* 'rowed' from *greb-* < \*g<sup>h</sup>reb<sup>h</sup>-, etc.), and Latin, where some perfect forms preserve the ablaut and \*-s- but not the original endings (*uēxī* 'I conveyed' from *ueh-* < \*weǵ<sup>h</sup>-, etc.). Some believe that the pl./du. originally had the \*e-grade (which would then be generalized in Greek and attested in the Vedic middle).

The thematic aorist has a zero-grade root (preserved in Vedic and Greek), thematic vowel (like in the present), and endings:

TABLE 1.39 IE THEMATIC AORIST

	PIE	Ved.	Gr.	OCS
1 sg.	*wid-o-m	á vidam <sup>b</sup>	ε íðon	<i>dvigъ</i>
2 sg.	*wid-e-s	á vidas	ε íðeç	<i>dviže</i>
3 sg.	*wid-e-t <sup>a</sup>	á vidat	ε íðe	<i>dviže</i>
1 pl.	*wid-o-me	á vidāma	ε íðome v	<i>dvigom {ъ}</i>
2 pl.	*wid-e-te	* á vidata	ε íðeete	<i>dvižete</i>
3 pl.	*wid-o-nt	á vidan	ε íðon	<i>dvigō</i>

<sup>a</sup> Cf. Arm. *egit* ‘found’; <sup>b</sup> cf. also augmentless *vidam*.

Cf. also Gr. ἤλυθε, OIr. *luid* < PIE \*h<sub>1</sub>lud<sup>h</sup>et ‘went’, which can hardly be secondary. For the zero grade, cf. also Gr. ἀφύγον, ἔτραπον from pres. φεύγω ‘I run’, τρέπω ‘I turn’, etc. Thematic aorists are often secondary, and exact correspondences are few – some researchers consider all (or most) of them as secondary/innovative. PIE may also have had a **reduplicated aorist**, cf. Ved. *ávocam*, Gr. εἶπον < \*(h<sub>1</sub>e)-we-wk<sup>w</sup>-o-m ‘I said’. The dual endings in Vedic/Greek differed from those of the present (the thematic vowel is in brackets): 1 \*(o)-we > Ved. -(ā)va, OCS -(o)v\ǎ; 2 \*(e)-to-m > Ved. -(a)tam, Gr. -(ε)tov; 3 \*(e)-te-h<sub>2</sub>m > Ved. -(a)tām, Gr. -(ε)την (< -(ε)τῆν).

Some also reconstruct **imperfect** for PIE, based on Greek and Vedic, cf. Ved. *ásam*, Gr. (Ion.) ἦα < (?) PIE \*(h<sub>1</sub>e)-h<sub>1</sub>es-ṃ; Ved. *ábharam*, Gr. ἔφερον < (?) PIE \*(h<sub>1</sub>e)-b<sup>h</sup>er-o-m. It is identical to the present except that it has aorist endings. This would then be the past tense of the imperfective aspect stem (the present being the present tense of imperfectives). Augmentless aorists and imperfects (sometimes used in a “timeless” meaning) found in Vedic, Gatha-Avestan, and Homeric/early Greek are often called **injunctives**. It seems that they were originally verbs unmarked for tense/mood when following preceding fully marked verbs or after the prohibitive imperative particle \*meh<sub>1</sub> ‘not’ (> Ved. *má*, Gr. μή, Alb. *mo* ‘do not!’, Arm. *mi*). In both cases, there was no need for redundant marking of tense/mood so it was left out, which is perhaps a structure of PIE origin – e.g., PIE may have used the injunctive forms like \*b<sup>h</sup>eront ‘they are carrying’ without the pres. ptc. \*-i, not the pres. \*b<sup>h</sup>eronti, after the present was already marked on a previous verb.<sup>12</sup> This is typologically very similar to the Bantu consecutive (cf. Nurse 2003: 101–102). For the middle aorist see p. 100.

## Perfect

The “perfect” originally (whether in the last phase of PIE or in pre-PIE) most likely indicated the stative aspect. The frequent, and plausible, assumption is that the original stative (expressing states like ‘I am dead’) at one stage, perhaps partially already in late PIE, became resultative (‘I have died [and I am dead now]’), and finally in post-PIE preterite (‘I died’). In most later IE languages, it has developed into a regular past tense. However, in Homeric Greek (with relics in Classical Greek) the perfect expresses the meaning of the state of the subject (this occurs less commonly also in R̥gveda), cf. Homeric τέθνηκε ‘is dead’, ἔσθηκε ‘is standing’ (cf. p. 309). Some original perfect forms have a present tense (or stative) meaning in other languages as well – like \*woydh<sub>2</sub>e ‘I know’ almost everywhere, Latin verbs like *memini* ‘I remember’ (cf. Goth. *man* ‘thinks’ and Gr. μέμνη ‘I yearn for’), or the Germanic *praeteritopraesentia* class (p. 400). The PIE perfect is

preserved in Indo-Iranian, Greek, Italic/Celtic (mixed with aorist), Germanic, and Anatolian (albeit in different form than in Indo-Iranian/Greek) – elsewhere, it was lost, usually with some remaining traces.

The usual perfect formation had partial reduplication of the root (originally perhaps only \*Ce- but also \*CV-, cf. Gr.  $\lambda\acute{\epsilon}\text{-}\lambda\omicron\upsilon\pi\alpha$  but Ved. *ri-réca* < \*le/i-loyk<sup>w</sup>-h<sub>2</sub>e from \*leyk<sup>w</sup>- ‘leave’), the ablaut alternation in the root (\*o-grade in the sg., zero grade in the pl. – preserved regularly in Vedic, sometimes in Germanic), and special (athematic) perfect endings (\*-h<sub>2</sub>e, \*-th<sub>2</sub>e, \*-e, \*-me, \*-e, \*-ī/-ēr). For reduplication cf. Ved. *ja-gama* ‘I went’ (from *gácchati* ‘goes’), Gr.  $\pi\acute{\epsilon}\text{-}\pi\omicron\iota\theta\alpha$  ‘I am persuaded’ (from  $\pi\epsilon\acute{\iota}\theta\omega$  ‘I persuade’), Lat. *ce-cinī* ‘I sang’ (from *canō* ‘I sing’), Goth. *hai-hait* ‘I called’ (from *haitan* ‘to call’), OIr. *ce-chan* ‘I sang’ (from *canid* ‘sings’), etc. Reduplication was regular in Indo-Iranian and Greek, and sporadic in Germanic, Italic, and Celtic. In PIE it may have been present in almost all forms (though we cannot be sure). The ablaut alternation is attested regularly in Indo-Iranian, in Germanic “strong” verbs, and some (early) Greek forms. The ablaut alternation includes the accent shift in Vedic (*jujósā* ‘he has enjoyed’ – *jujusúr* ‘they have enjoyed’) and Germanic “strong” verbs, cf. OEng. *cēas* ‘I/he chose’ – *curon* ‘we chose’ (with Verner’s Law – p. 20, 393) < PIE \*(ǵu)ǵowse – \*(ǵu)ǵusr. Anatolian shows no special perfect stem (cf. p. 187 for possible reduplication), just different endings. The exact relation of the PIE perfect to the Anatolian *hi*-present, in spite of the formally close similarity of endings, is controversial (p. 187).

The best-attested perfect form is the one that is rather aberrant – the perfect form \*woyd-/wid- has no reduplication (it is the only such plausibly reconstructable form) and has a stative (in later languages present) meaning ‘know’, derived from \*weyd- ‘see’ (p. 49). The apparent (already PIE) development was ‘I have seen’ > ‘I know’ (or, more precisely, ‘I am in the state of having seen/knowing’):

TABLE 1.40 IE PERFECT

	PIE	Ved.	Gr.	Goth.
1 sg.	*woyd-h <sub>2</sub> e <sup>a</sup>	<i>véda</i>	οἶδα	<i>wait</i>
2 sg.	*woyd-th <sub>2</sub> e <sup>b</sup>	<i>véttha</i>	οἶσθα	<i>waist</i>
3 sg.	*woyd-e	<i>véda</i>	οἶδε	<i>wait</i>
1 pl.	*wid-me	<i>vidmá</i>	ἴδμε v	<i>wit u m</i>
2 pl.	*wid-e	<i>vidá</i>	ἴσ τε	<i>wit u þ</i>
3 pl.	*wid-ī/-ēr	<i>vidúr</i>	ἴσ {α σ }	<i>wit un </i>
1 du.	*wid-we	<i>vidvá</i>		<i>witu</i>
2 du.	*wid-tH-(?) <sup>c</sup>	<i>vid a th úr</i>		<i>wit u ts</i>
3 du.	*wid-t-(?) <sup>c</sup>	<i>vid a t úr</i>		

<sup>a</sup> Cf. OCS isolated *vědě* ‘I know’ < \*woydh<sub>2</sub>e-y (with *hic* & *nunc* \*-i); <sup>b</sup> phonetically \*woysth<sub>2</sub>e (p. 54), cf. Hitt. pret. -*ta*; <sup>c</sup> phonetically \*witst(H)- (p. 54).

The PIE perfect endings were close to both the pres./aor. active and middle endings, though the latter is often overly emphasized. The 1 sg. \*-h<sub>2</sub>e is attested directly as *-h̄ha* in Anatolian (p. 187), and indirectly in the lack of Brugmann’s Law (p. 40, 206) in Ved. *cakara* ‘I have made’ < \*k<sup>w</sup>e-k<sup>w</sup>or-h<sub>2</sub>e (a closed syllable because of the laryngeal) but *cakára* ‘he has made’ < \*k<sup>w</sup>e-k<sup>w</sup>or-e (an open syllable). The opposition of the 1 sg. \*-a < \*-h<sub>2</sub>e and the 3 sg. \*-e is seen in the OIr. pret. 1 sg. *gád* ‘I asked’ but 3 sg. *gáid* ‘he asked’ (with palatalization form \*-e). The 1 pl. \*-me and the 1 du. \*-we is the same as in the pres./aor.; the 2 pl. \*-e (\*-o is theoretically also possible) is attested only in Vedic but looks archaic (elsewhere we find the pres./aor. \*-te). The variants in the 3 pl. are not clear – cf. Hitt. pret.

-*ēr*, older Lat. -*ēr-e* (with a secondary particle \*-i) for \*-*ēr* (from original \*-*er-s*?, cf. p. 70); Av. -*ar*<sup>o</sup>, probably Ved. -*ur*, and Hitt. rare -*ar* for \*-*ṛ* (cf. also NPhryg. δακαρ(εν) ‘they have made’); (?) GAv. *cikōitər*<sup>o</sup>*š* (and perhaps Ved. -*ur* and Lyd. -*rš*) for \*-*ṛ-s* (cf. also Ven. -*ers* (?)). The particle (?) \*-*s* in some 3 pl. forms is perhaps the same as in the present \*-*me/o-s* (and possibly identical to the nominal nom. pl. \*-*s*?). The original perfect endings have been influenced by the aorist ones in a number of languages. The intrusion of the present \*-i (p. 92–93) into Italic perfect endings (cf. Lat. 1 sg. -*ī* < early -*ei* < \*-*h<sub>2</sub>e-y*, 2 sg. -*is-tī* < \*-*th<sub>2</sub>e-y*) points to the older stative (“present”) meaning of the perfect.

**Middle (mediopassive)**

The middle (medium) voice stands in opposition to the active voice (the present/aorist forms we have seen were all active – the perfect had no active/middle distinction but can be considered active of a sort). The reconstruction of the concrete meaning of the definitely poly-functional PIE middle is rather speculative. The active typically expressed a simple action (or state) by the subject (cf. Gr. λύω ‘I unbind (something)’), while the middle indicated that the subject is doing an action for his or her benefit or that it is affected by the action (cf. Gr. λύομαι ‘I unbind (for) myself’). The meaning was thus reflexive (Gr. λούομαι ‘I wash myself’) and sometimes reciprocal (cf. Ved. *vádetē* ‘two of them talk to each other’). Since the PIE middle yielded a passive voice (cf. Goth. *gibada* ‘is given’) in some families (Italic, Germanic), it seems that the PIE middle also had that function in some cases (hence the other name, mediopassive). It also may have had an impersonal function (cf. Lat. *itur* ‘one goes’). Some verbs (called deponent or *media tantum*) had exclusively middle endings but non-middle meaning (cf. Ved. *śáye*, CLuw. *zīyari* < PIE \**key-or* ‘lies’; Ved. *váste*, Hitt. *wešta* < PIE \**wes-tor* ‘wears’). Most families preserved the PIE middle (the opposition active : middle is best preserved in Anatolian, Indo-Iranian, and Greek), with the notable exceptions of Balto-Slavic (which preserves traces in active endings, cf. OPruss. *assai* ‘you are’ < \*-*so-y*) and Armenian (with innovative mediopassive forms).

The middle had special endings – for the middle present and middle aorist – similar to both the present/aorist and perfect endings (though the latter is usually overemphasized), the reconstruction of which is rather difficult. It was typified by two particles: \*-*o* (in both the present and aorist) and \*-*r* (only in the present). Instead of \*-*r* (attested in Anatolian, Italian, Celtic, Armenian, Tocharian, Phrygian, Venetic), posited as original by many researchers, some dialects have the *hic & nunc* \*-*y* (Indo-Iranian, Greek, Germanic, Balto-Slavic, Albanian), which is likely secondary but possibly a dialectal feature already in PIE. In the 2/3 sg. and 3 pl., variant endings can be reconstructed (present/aorist-like and perfect-like) – their original function/distribution is not clear. Here is a tentative reconstruction of the present (primary) middle endings (both thematic and athematic):

**TABLE 1.41 IE MIDDLE PRESENT**

	PIE	Ved.	Gr.	Lat.	Goth.	OIr.	Toch. B	Hitt.
1 sg.	*- <i>h<sub>2</sub>e-(r)</i>	- <i>ē</i> \	-  <i>u σ t </i>	{- <i>or</i> }	{- <i>da</i> } <sup>f</sup>	{- <i>ur</i> }	{- <i>mar</i> }	-( <i>h</i> ) <i>ḫa(r ī)</i>
2 sg.	*- <i>s-o</i>	- <i>sē</i> \	- <i>σ t </i>	- <i>re</i>	- <i>zā</i> \			
	*- <i>th<sub>2</sub>e-r</i>		(dial.) <sup>d</sup>			- <i>thē</i> <i>r</i>	- <i>ta</i> <i>r</i> \	- <i>ta(r ī)</i>
3 sg.	*- <i>t-o-r</i> <sup>a</sup>	- <i>tē</i> \	- <i>τ t </i>	- <i>tur</i>	- <i>dā</i> \	- <i>thar</i>	{- <i>tār</i> }	- <i>ta(r ī)</i>
	*- <i>o-r</i>	- <i>e</i> \	(dial.) <sup>e</sup>			- <i>ar</i>		- <i>a(r ī)</i>
1 pl.	*- <i>me-(s)-d<sup>h</sup>h<sub>2</sub></i>	- <i>mā</i> <i>hē</i> \	- <i>μ(σ)θ α</i>	{- <i>mur</i> }	- <i>ndā</i> \	{- <i>mmar</i> }	- <i>mtā</i> <i>r</i>   <i>r</i>	- <i>w<sup>h</sup>ašta</i>
2 pl.	*- <i>(s)-d<sup>h</sup>we</i>	- <i>dhvā</i> <i>e</i> \	- <i>σθ ε</i>	{- <i>minī</i> }	{- <i>nda</i> }	- <i>id</i>	- <i>tā</i> <i>r</i>   <i>r</i>	- <i>tum<sup>h</sup>a</i> <sup>g</sup>
3 pl.	*- <i>nt-o-r</i>	- <i>ntē</i> \	- <i>ντ t </i>	- <i>ntur</i>	- <i>ndā</i> \	- <i>tar</i>	{- <i>ntār</i> }	- <i>anta(r ī)</i>
	*-( <i>ē</i> ) <i>r-o</i> <sup>b</sup>	- <i>rē</i> \	(dial.)					



<sup>a</sup> Cf. Phryg. -top; <sup>b</sup> also variant \*-(ē)r-o-r?; <sup>c</sup> cf. Av. *āṅhārē* ‘they are sitting’ < \*āsārai < \*-ēr-oy; <sup>d</sup> Arc. κει-ου ‘you lie’; <sup>e</sup> Cyp. κει-του ‘he lies’; <sup>f</sup> cf. Oic. *heite* ‘I am called’ < \*-ai < PIE dial. \*-oy; <sup>g</sup> cf. Luw. *-tuwari*

For the 1 sg. (best preserved in Anatolian) cf. thematic pres. \*-h<sub>2</sub> and perf. \*-h<sub>2</sub>e (some reconstruct \*-h<sub>2</sub>o(r) for the middle); for the 2 sg. pres./aor. \*-s- and perf. \*-th<sub>2</sub>e (some reconstruct \*-th<sub>2</sub>o(r) for the middle); for the 3 sg. pres./aor. athematic \*-t- and thematic pres./perf. \*-e; for the 3 pl. pres./aor. \*-nt- and perf. \*-t/-ēr (the \*n/r here reminds one of heteroclitic nouns – p. 77–78). Some linguists reconstruct the PIE forms without the particle \*-r, which they believe to be secondary (stemming supposedly, not too convincingly, from the 3 pl. ending). However, it is not impossible that \*-r was facultative or dialectal in PIE. The present/aorist-like endings \*-so, \*-to(r), \*-nto(r) are sometimes thought to be innovative (with \*-o originating perhaps in the 3 sg. ending \*-o), but they must have been there already in PIE. Others think that variant endings point to two different original categories<sup>13</sup> (one with the ending \*-to and one with the ending \*-o, etc.). Ved. forms like *brūtē* ‘calls (for himself)’ but *bruve* (pass.) ‘is called’ and some other facts would perhaps point to a possible original meaning distinction. In the 1 pl., \*-me- is the ending attested in all verbal finite forms (the optional \*(s)-, attested in Greek/Hittite, is perhaps related to the pres. \*-mes and then transferred to 2 pl. \*(s)-d<sup>h</sup>we), and \*-d<sup>h</sup>h<sub>2</sub> is probably some kind of particle. As in the active, the -w- in Hitt. *-wašta* may be from the old du. \*-we-(s)-d<sup>h</sup>h<sub>2</sub> (cf. Ved. 1 du. pres. *-vahe*, aor. *-vahi*, Av. aor. *-vadi*). The 2 pl. \*-d<sup>h</sup>we is unique to the middle, though the final \*-e (cf. Gr. -σθε), perhaps accidentally, reminds one of the pres./aor. \*-te and perf. \*-e; some, however, reconstruct \*-d<sup>h</sup>wo (cf. Luw. *-tuwa-ri*). Many later endings (like the 1 pl. Lat. *-mur*, OIr. *-mmar*) are remodeled by analogy to the active. A thematic *media tantum* present with many reflexes is PIE \*sek<sup>w</sup>-e-tor ‘follows’ > Ved. *sácate*, Gr. *ἕπεται*, Lat. *sequitur*, OIr. *sechithir*.

These are the aorist (“secondary”) endings:

TABLE 1.42 IE MIDDLE AORIST

	PIE	Ved.	Gr.	Toch. B	Hitt.
1 sg.	*-h <sub>2</sub>	-i	{-μην}	{-mai}	-ḫ at(i)
2 sg.	*-s-o	<sup>a</sup>	-σο		
	*-th <sub>2</sub> e	-th <sup>h</sup> ā s		-ta i	-ta
3 sg.	*-t-o	-ta	-το	-te	-ta
	*-o	-a			-a t(i)
1 pl.	*-me-(s)-d <sup>h</sup> h <sub>2</sub>	-mahi	-με(σ)θα	-mte	-wašta t
2 pl.	*-(s)-d <sup>h</sup> we	-dhva m	-σθε	-t	-dum <sup>h</sup> a t
3 pl.	*-nt-o	-ata, -nta	-ντο	-nte	-anta t(i)
	*-r-o	-ra n			

<sup>a</sup> Cf. Av. *-sa*.

The endings were probably mostly the same as the present ones but with no \*-r (i.e., dial. \*-y). Ved./Av. *-i* would point to \*-h<sub>2</sub> (unlike the present \*-h<sub>2</sub>e(r)). Perhaps the 1 pl. \*-mes-d<sup>h</sup>h<sub>2</sub> was originally the present ending and \*-me-d<sup>h</sup>h<sub>2</sub> the aorist ending.

## Subjunctive

The present/aorist forms we have seen were all indicative (the perfect had no mood distinction but can be considered indicative of a sort). The subjunctive and optative were originally probably normal verbal suffixes (like many others that existed in the present – p. 94–95) that only later became full-fledged moods. The PIE subjunctive probably expressed something that the speaker thought was uncertain, a future action with a slight

reservation (e.g., \*h<sub>1</sub>esoh<sub>2</sub> ‘I suppose/guess I will be’). Thus, it is understandable that it yielded the future tense in Latin. The formation of the subjunctive was simple – root in the full grade + suffix \*-e-/o- + present/aorist endings. The suffix \*-e-/o- was formally (and probably etymologically) identical to the thematic vowel and followed the same qualitative pattern. The subjunctive thus in a way presents the “thematized” forms of the indicative, where the already thematic forms become “doubly thematic” in the subjunctive (with \*-e-e-, \*-o-o-) and the athematic forms become thematic. Why the thematic vowel (or a morpheme formally identical to it) functions as a subjunctive suffix is not clear. These are the forms of athematic verbs:

**TABLE 1.43 IE ATHEMATIC SUBJUNCTIVE**

	PIE	Ved.	Gr. (Hom.)	Lat. <sup>e</sup>
1 sg.	*h <sub>1</sub> es-o-h <sub>2</sub>	<i>ásā ni </i> <sup>c</sup>	ἔω	<i>erō</i>
2 sg.	*h <sub>1</sub> es-e-s-(i) <sup>a</sup>	<i>ásas(i)</i>	ἔῃ ς	<i>eris</i>
3 sg.	*h <sub>1</sub> es-e-t-(i) <sup>b</sup>	<i>ásat(i)</i>	ἔῃ	<i>erit</i>
1 pl.	*h <sub>1</sub> es-o-me/(o-(s/n))	<i>ásāma</i>	ἔω μεν <sup>d</sup>	<i>erimus</i>
2 pl.	*h <sub>1</sub> es-e-te	<i>ásat h a</i>	ἔῃ τε	<i>eritis </i>
3 pl.	*h <sub>1</sub> es-o-nt-(i)	<i>ásan</i>	ἔω σι	<i>erunt</i>

<sup>a</sup> Also \*-e-h<sub>1</sub>-i? (p. 93–94); <sup>b</sup> also \*-e-e? (p. 93–94); <sup>c</sup> cf. the older ending in *brav-ā* ‘I speak’ (-ā 13x in RV); <sup>d</sup> cf. Hom. ἴ-ο-μεν ‘we would go’ for the original short \*-o-; <sup>e</sup> future tense.

It is not clear whether the endings in PIE were originally primary, secondary, or both (either mixed or facultative). Vedic has a curious mix of primary and secondary endings. The 1 sg. shows only (thematic primary/present) \*-h<sub>2</sub> (unlike in the optative), which is in accord with the fact that all subjunctive forms were “thematic”. Thus, other primary endings might be expected as well. However, some Vedic endings (1 pl. *-ma*, 3 pl. *-n*, 1 du. *-va*), parallelism to the optative, and structural reasons (a mood unmarked for present/aorist could have used generic, and not present, endings) could point to originally secondary endings. Both primary and secondary endings can be secondary in certain cases in later languages. If variants existed already in PIE, they probably carried no special (present/aorist) semantics. The forms of thematic verbs are:

**TABLE 1.44 IE THEMATIC SUBJUNCTIVE**

	PIE	Ved.	Gr.	Lat. <sup>f</sup>
1 sg.	*b <sup>h</sup> er-o-o-h <sub>2</sub>	<i>váhā ni </i> <sup>c</sup>	φέρω	<i>fer{am}</i>
2 sg.	*b <sup>h</sup> er-e-e-s-(i) <sup>a</sup>	<i>bhárās</i> <sup>d</sup>	φέρῃς	<i>ferēs</i>
3 sg.	*b <sup>h</sup> er-e-e-t-(i) <sup>b</sup>	<i>bhárāt(i)</i>	φέρῃ <sup>e</sup>	<i>feret</i> <sup>g</sup>
1 pl.	*b <sup>h</sup> er-o-o-me/(o-(s/n))	<i>bhárāma</i>	φέρωμεν	<i>ferē mus</i>
2 pl.	*b <sup>h</sup> er-e-e-te	<i>váhāt h a</i>	φέρῃτε	<i>ferē tis </i>
3 pl.	*b <sup>h</sup> er-o-o-nt-(i)	<i>vahān</i>	φέρωσι	<i>ferē nt</i>

<sup>a</sup> Also \*-e-e-h<sub>1</sub>-i? (p. 93–94); <sup>b</sup> also \*-e-e? (p. 93–94); <sup>c</sup> cf. older *arc-ā* ‘I may shine/praise’, GAv. *yaojā* ‘I will yoke’; <sup>d</sup> cf. *vahāsi*; <sup>e</sup> cf. Arc. εῤῥῃ ‘has’, Thess. θεῤῥῃ ‘wishes’ < \*-ēt (?); <sup>f</sup> future tense (3rd/4th conjugation); <sup>g</sup> with *-t* < \**ti*, cf. Osc. *fakiiad* ‘makes’ < \**t*.

\*-ee- contracted to \*-ē-, and \*-o-o- to \*-ō- (others believe that It. *-ē-* in all persons is actually archaic). The contraction was probably post-PIE (Gatha-Avestan still has two syllables there).

## Optative

The optative probably expressed the wishes of the speaker (e.g., \**h<sub>1</sub>syeh<sub>1</sub>m* ‘I would be’). An athematic and thematic optative existed, though the latter was perhaps very late PIE, post-Tocharian (p. 91–92, 455). The PIE optative yielded the (so-called) subjunctive mood in Germanic and Italic, and the imperative in Balto-Slavic. The athematic forms were made of the zero-grade root, an alternating suffix (sg. \*-yeh<sub>1</sub>-, pl. \*-ih<sub>1</sub>-), and mostly generic (“aorist/secondary”) endings:

TABLE 1.45 IE ATHEMATIC OPTATIVE

	PIE	Ved.	Gr. (Hom.)	OLat. <sup>c</sup>
1 sg.	* <i>h<sub>1</sub>s-yeh<sub>1</sub>-m</i>	<i>syám</i>	εἶην	<i>siem</i>
2 sg.	* <i>h<sub>1</sub>s-yeh<sub>1</sub>-s</i> <sup>a</sup>	<i>syás</i>	εἶης	<i>siē̃s</i> <sup>\</sup>
3 sg.	* <i>h<sub>1</sub>s-yeh<sub>1</sub>-t</i> <sup>b</sup>	<i>syát</i>	εἶη	<i>siē̃</i> <sup>\</sup> <sup>d</sup>
1 pl.	* <i>h<sub>1</sub>s-ih<sub>1</sub>-me</i>	<i>sy<sup>\</sup>á̃ma</i>	εἶμε	<i>sīm</i> { <i>us</i> }
2 pl.	* <i>h<sub>1</sub>s-ih<sub>1</sub>-te</i>	<i>sy<sup>\</sup>á̃ta</i>	εἶτε	<i>sīt</i> { <i>is</i> }
3 pl.	* <i>h<sub>1</sub>s-ih<sub>1</sub>-ent</i>	<i>sy<sup>\</sup>úr</i> }	εἶεν	<i>sien</i> <sup>\</sup>

<sup>a</sup> Cf. the OCS athematic imp. *ěžd-b!* ‘eat!’ < \**h<sub>1</sub>ed-y-* (also *dažd-b!* ‘give!’); <sup>b</sup> cf. OCS *eša* ‘if only’ < \**h<sub>1</sub>esyeh<sub>1</sub>t* (with secondary \**h<sub>1</sub>es-*); <sup>c</sup> subjunctive; <sup>d</sup> with secondary *-t* < \**-ti*, cf. oldest SIED with *-d* < \**-t*.

The ablaut scheme, typical for athematic paradigms, is preserved in Greek and Latin. The zero-grade \*-ih<sub>1</sub>- can also be seen in Vedic (middle opt. *bruv-ī-* ‘talk’), Old Church Slavic (imp. *dad-i-te!* ‘give (pl.)!’), Lithuanian (dial. imp. *sėdžīte!* ‘remain (pl.) seated!’), Gothic (pres. *wil-ei-ma* ‘we will’, cf. Lat. subj. *uel-ī-mus* ‘we may want’), and Tocharian (opt. *klyauš-i-m* ‘I might hear’, p. 468). The 3 pl. ending was \*-ent from the athematic present (not the aorist \*-nt), based on Latin/Greek. The thematic optative had a full-grade root + thematic vowel \*-o- + suffix \*-ih<sub>1</sub>- + generic (“aorist”) endings:

TABLE 1.46 IE THEMATIC OPTATIVE

	PIE	Ved.	Gr.	OCS <sup>d</sup>	Goth.
1 sg.	* <i>b<sup>h</sup>er-o-yh<sub>1</sub>-m</i>	<i>bhav<sup>\</sup>e ya m</i> <sup>a</sup>	φέροι{ <i>μι</i> } <sup>c</sup>	<i>bōdē</i> { <i>mь</i> } <sup>e</sup>	<i>bairau</i>
2 sg.	* <i>b<sup>h</sup>er-o-yh<sub>1</sub>-s</i>	<i>bhaves</i> <sup>a</sup>	φέροις	<i>beri!</i> <sup>f</sup>	<i>bairais</i>
3 sg.	* <i>b<sup>h</sup>er-o-yh<sub>1</sub>-t</i>	<i>bhāvet</i> <sup>a</sup>	φέροι	<i>beri!</i>	<i>bairai</i>
1 pl.	* <i>b<sup>h</sup>er-o-yh<sub>1</sub>-me</i>	<i>bharema</i>	φέροιμε	<i>berēm</i> <sup>\</sup> <sup>!</sup>	<i>bairaim</i> <sup>\</sup> <sup>!</sup>
2 pl.	* <i>b<sup>h</sup>er-o-yh<sub>1</sub>-te</i>	Skr. <i>bhareta</i>	φέροιτε	<i>berēte!</i> <sup>g</sup>	<i>bairaiþ</i>
3 pl.	* <i>b<sup>h</sup>er-o-yh<sub>1</sub>-ent</i>	<i>vah<sup>\</sup>e y ur</i> } <sup>b</sup>	φέροιεν		<i>bairain</i>   <i>a</i>

<sup>a</sup> \**b<sup>h</sup>uh<sub>2</sub>-* ‘be’; <sup>b</sup> \**weg<sup>h</sup>-* ‘convey’, cf. the old ending in Av. *baraiīn*; <sup>c</sup> cf. Arc. *εἴξελαινοῦα* ‘I would drive out’; <sup>d</sup> imperative; <sup>e</sup> ‘I would be’ (\**b<sup>h</sup>uh<sub>2</sub>-* ‘be’); <sup>f</sup> cf. OPruss. *wedais!* ‘lead (sg.)!’; <sup>g</sup> cf. OLith./dial. *sākaite!* ‘will you (pl.) please say!’

The lack of qualitative ablaut of the thematic vowel is conspicuous. In most languages, the suffix \*-oyh<sub>1</sub>- yields (or reflects just like) \*-oy-. However, the Slavic acute in \*-ēme/o/ь, \*-ēte would point to a laryngeal.

## Imperative

The imperative, easily reconstructable, had forms only for the 2nd and 3rd persons. It had eventive endings in the 3 sg. and 2/3 pl., and was characterized by the particle \*-u in the 3 sg./pl. The athematic forms were:

TABLE 1.47 IE ATHEMATIC IMPERATIVE

	PIE	Ved.	Gr.	Lat.	Hitt.
2 sg.	*h <sub>1</sub> ey *h <sub>1</sub> i-d <sup>h</sup> i <sup>a</sup>	<i>ihī</i> <sup>b</sup>	ἴθι	<i>ī</i>	<i>ēp</i> <sup>d</sup> <i>īi</i>
3 sg.	*h <sub>1</sub> ey-t-u	<i>etu</i> <sup>c</sup>			<i>ēptu</i> <sup>d</sup>
2 pl.	*h <sub>1</sub> i-te	<i>itá</i>	ἴτε	<i>īte</i>	<i>īite n </i>
3 pl.	*h <sub>1</sub> y-ent-u	<i>yántu</i>			<i>app a vntu</i> <sup>d</sup>

<sup>a</sup> Cf. OCS *i-dā!* ‘go!’, *věžd-ō!* ‘know!’; <sup>b</sup> cf. *śru-dhi!* ‘hear!’ for *-dh-*; <sup>c</sup> cf. *ástu* < \*h<sub>1</sub>es-tu for the accent; <sup>d</sup> *epp-* ‘take’ (PIE \*h<sub>1</sub>ep-).

The 2 sg. is either the full-grade endingless root or the zero-grade root plus \*-d<sup>h</sup>i (probably a particle). Except for the \*-d<sup>h</sup>i form, the singular has the full-grade and the plural the zero-grade root, as usual in the athematic stems. The thematic forms were:

TABLE 1.48 IE THEMATIC IMPERATIVE

	PIE	Ved.	Gr.	Lat.	Goth.
2 sg.	*b <sup>h</sup> er-e	<i>bhára</i>	φέρε	<i>age</i> <sup>a</sup>	<i>bair</i>
3 sg.	*b <sup>h</sup> er-e-t-u	<i>bháratu</i>			<i>bair a d au </i> <sup>b</sup>
2 pl.	*b <sup>h</sup> er-e-te	<i>bhárata</i>	φέρετε	<i>agite</i>	<i>bairiþ</i>
3 pl.	*b <sup>h</sup> er-o-nt-u	<i>bhárantu</i>			<i>bair adau </i>

<sup>a</sup> *ag-* ‘drive’ (PIE \*h<sub>2</sub>eǵ-); <sup>b</sup> as if from a full-grade \*-t-ow?

The 2 sg. has a zero-ending (cf. also OIr. *beir!*). The 2 pl. imp. seems to have been identical in some cases (like \*h<sub>1</sub>ite, \*b<sup>h</sup>erete) to the 2 pl. pres./inj.

There were also special future imperative forms that express the meaning of a delayed (future) command (\*h<sub>1</sub>itōt ‘go later!’):

TABLE 1.49 IE FUTURE IMPERATIVE

	PIE	Ved. <sup>a</sup>	Gr. <sup>c</sup>	Lat.
2/3 sg., 2 pl.	*h <sub>1</sub> i-tōt	<i>vittāt</i> <sup>b</sup>	ἴτω	<i>itō</i> <sup>d</sup>
2/3 sg., 2 pl.	*b <sup>h</sup> er-e-tōt		φέρετω	<i>agitō</i>
3 pl.	*-(o)-n(t)-tōt		dial. -vτω	<i>euntō, aguntō</i>

<sup>a</sup> Usually just 2 sg.; <sup>b</sup> ‘thou shalt regard’; <sup>c</sup> used as regular imperative for the 3 sg.; <sup>d</sup> cf. arch. Lat. *da-tōd* ‘let him give’, Clb. *ta-tuz* ‘he must give’; only 2/3 sg. in Latin.

The ending \*-tōt (often reconstructed as \*-tōd) is usually explained as originally an abl. sg. (p. 63) of the pronoun \*tod ‘that’, meaning originally ‘from that > after that’ or

the like, but the pronominal form was actually \**tosmōt* (p. 85). The 2/3 sg. and 2 pl. had the same ending (probably from older \*-*Ø-tōt*, \*-*et-tōt*, \*-*ete-tōt*), and the 3 pl. was made from \*-*nt-* + \*-*tōt-*.

### Non-finite verb forms

Unlike many later IE languages, PIE had no infinitive but had participles, verbal adjectives, and verbal nouns. The participles were those of the active present, middle, and perfect. The present active participle had the \*-*ont-* suffix (with no thematic vowel before it): m.  $\pm\pm\pm$  holodynamic \**b<sup>h</sup>er-ōnt-s* ‘carrying’ (Ved. *bhár<sup>h</sup>a<sup>h</sup>ṇ*, Gr. *φέρων*, Lat. *fer<sup>h</sup>ēns*, OCS *bery*, Goth. *baír<sup>h</sup>a<sup>h</sup>nds*; Lith. *neš<sup>h</sup>ā<sup>h</sup>s*; cf. Hitt. *ašanza* ‘being’ (active!) but *kunanza* ‘killed’) – gen. sg. \**b<sup>h</sup>er-nt-e/os* (Ved. *bháratas*, Lat. *ferentis*); f. \**b<sup>h</sup>er-ont-ih<sub>2</sub>* (Ved. *bhárantī*, Gr. *φέρουσα* – cf. p. 38 for “laryngeal breaking”, OCS *berq<sup>h</sup>št<sup>h</sup>l<sup>h</sup>*; Lith. *nėšanti*) – gen. sg. \**b<sup>h</sup>er-ont-yeh<sub>2</sub>-es* (OCS *berq<sup>h</sup>št<sup>h</sup>{e}*), Lith. *nėšančios*) (cf. p. 67). Cf. also the athematic stem \**h<sub>1</sub>s-ōnt-s* ‘being’ (Lat. *sōns* ‘guilty’ < \*‘he who is’, OCS *sy*), \**h<sub>1</sub>s-nt-ih<sub>2</sub>* (f.) (Ved. *satī*, Gr. Dor. *ἔασσα*). The present middle participle was a simple *o*-stem: \**b<sup>h</sup>er-o-mh<sub>1</sub>n-o-s* ‘carrying for himself/carried’ (Ved. *bhāram<sup>h</sup>ā<sup>h</sup>nas*, YAv. *barəmna-*, Gr. *φέρόμενος*; OCS *nesom<sup>h</sup>*, Lith. *nėšamas* ‘being carried’; CLuw. *kīšama-* ‘combed’; cf. Lat. *alumnus* ‘nursling’) – f. \**mh<sub>1</sub>n-e-h<sub>2</sub>* (cf. Lat. *fēmīna* ‘woman’ < \**d<sup>h</sup>eh<sub>1</sub>mh<sub>1</sub>neh<sub>2</sub>* ‘she that is sucked’); athematic \*-*ṃh<sub>1</sub>nos* (Ved. *vidāná-vidāna-* ‘finding’) The perfect participle was an *s*-stem (\*-*wos-* suffix): m.  $\pm\pm\pm$  holodynamic \**weyd-wōs* ‘in the state of knowing’ (Ved. *vā<sup>h</sup>dvā<sup>h</sup>ṃs-*, GAv. *vā<sup>h</sup>dvuā*, Gr. *εἰδώς* ‘knowing’; Goth. *weitwo<sup>h</sup>d<sup>h</sup>s* ‘witness’) – oblique \**wid-us-* (Ved. *viduṣ-*, OCS *nes<sup>h</sup>š-* ‘having carried’, Lith. *likus-* ‘having left’, cf. Goth. noun *berusjos* ‘parents’); f. \**wid-us-ih<sub>2</sub>* (Ved. *vidúṣī*, Gr. Hom. *ιδυῖα*, OCS *nes<sup>h</sup>ši*, Lith. *likusi*), n. \**weyd-wos* (Gr. *εἰδός*). Unlike the participle of \**woyd<sub>2</sub>e* (p. 98), other perfect participles were originally reduplicated (cf. Ved. *ja-ghan-vā<sup>h</sup>ṃs-* from PIE \**g<sup>wh</sup>en-* ‘beat’).

Besides the three participles, embedded paradigmatically in the PIE verbal system, there were also some verbal adjectives and nouns, less directly connected (in the realm of word formation) to verbal finite forms. The zero-grade verbal adjective in \*-*to-* is widely attested, e.g., \**sth<sub>2</sub>-t-o-s* ‘placed’ (Ved. *sthitás* ‘standing, settled’, Gr. *στατός* ‘placed, standing’, Lat. *status* ‘set’; cf. OCS *pro-str<sup>h</sup>ṛṛṛ* ‘spread’, Goth. *nasips* ‘saved’). The adjectives in \*-*n-o-s* (cf. Ved. *vinnás* besides *vittás* ‘found’; OCS *znan<sup>h</sup>*, OEng. *ge-cnāwen* ‘known’; cf. also \**p<sup>h</sup>h<sub>1</sub>nos* – p. 37) of a similar function are not as widely attested. The suffix \*-*t-* was present also in the *i*-stem verbal noun: \**g<sup>w</sup>ṃ-t-i-s* ‘going’ (Ved. *gátis*, Gr. *βάσις* ‘stepping’). The acc. sg. of the \*-*t-u-s* nouns might have been used as a supine (after motion verbs): \*-*t-u-m* (cf. Lat. sup. *-tum*, OCS sup. *-ṛṛ*, Lith. dial. sup. *-ty*, Ved. inf. *-tum*), cf. Lat. *cubi-tum uēnerunt* ‘they went to sleep’, OCS *id<sup>h</sup> lovi-ṛṛ* ‘I go to hunt’.

### POSTPOSITIONS

A considerable number of postpositions can be reconstructed for PIE, with some of them functioning also as adverbs (p. 80–81). Most modern and some older IE languages (like Latin, Greek, Old Church Slavonic, or Gothic) predominantly have prepositions (which are to have an even more important role in younger IE languages with simplified or lost declensions), which often function as verbal prefixes (modifying the meaning of the verb) as well, cf. the Lat. preposition *ad* ‘to, toward’ and *ad-ueniō* ‘I come to’. However, these verbal prefixes were originally independent of verbs (cf. Eng. *come to*), and some more archaic old IE languages (Anatolian, Vedic) predominantly had postpositions, which is

mostly held to have been the case in PIE as well (though it is possible that there were also some prepositions, or that at least some postpositions could have also been placed prepositionally, etc.). Some postpositions can be found in languages that are mostly prepositional (like Latin, Greek, Old Persian), cf. Lat. *mē-cum* ‘with me [lit. me-with]’. As examples of reconstructed postpositions (later becoming prepositions in many languages) cf., e.g., \*(h<sub>1</sub>)en ‘in’ (Gr. ἐν, Lat. *in*, OPruss. *en*, Goth./Eng. *in*, OIr. *i*), \*h<sub>2</sub>epo ‘(away) from’ (Ved. *āpa*, Gr. ἀπό/ἄπο, Lat. *ab*, Goth. *af*, Eng. *of*), etc. There are a few suffixes that appear in postpositions like \*-ti, \*-b<sup>hi</sup>, \*-r(i), cf. \*pro-ti ‘towards’ (Ved. *prāti*, Gr. epic προτί, OCS *prot-ivъ* ‘against’ – Slav. also \*protъ) from \*prō (Ved. *prá*, Gr. πρό ‘before, forth’, Lat. *prō* ‘in front of, before’, OCS *pro-*, Lith. *pra-* ‘through’, Goth. *fra-*, Eng. *from*, OIr. *ro-*). As already said, PIE (mostly plural) case markers, like instr. \*-b<sup>hi</sup>, were also originally postpositions, and some of them were not completely transformed to proper endings (p. 63).

## CONJUNCTIONS

Not a lot of conjunctions can be reconstructed. It is an old idea that PIE had almost no subordinate clauses (or that they were at least rare)<sup>14</sup> and expressed those constructions mainly through participles (e.g., ‘the going man’ instead of ‘the man that is going’ – but cf. the relative pronoun \*yo-, p. 88–89), verbal nouns, and likely pronominal forms used as conjunctions. This means that the syntax of the subordinate clauses (and its conjunctions) is mostly innovative in later IE languages. However, not even the conjunctions in independent clauses were stable. The two best-known independent-clause reconstructed conjunctions are \*-k<sup>we</sup> ‘and’ (Ved. *ca*, GA. *-cā*, Gr. τε, Lat. *-que*, Clb. *-kue*, OIr. *-ch*, W *-p*, Goth. *-h*, Hitt. *-kku*; for the negative \*nek<sup>we</sup> cf. below) and \*-wě (with optional monosyllabic lengthening – p. 54–55) ‘or’ (Ved./GA. *vā*, Lat. *-ue*, Clb. *-ue*, Toch. B *wa-t*). Unlike ‘and’ and ‘or’ in modern IE languages, these were enclitics (placed after the second word – one could call them postpositive particles), cf. Lat. *Senātus populus-que Rōmānus* ‘the Senate and people [lit. people-and] of Rome’, Ved. *Mitrām huve Vārunam ca* ‘I invoke Mitra and Varuna [lit. Varuna-and]’. As in modern languages, they could be used twice with slightly different meanings (\*w<sup>l</sup>k<sup>w</sup>ōs h<sub>2</sub>ekwōs-k<sup>we</sup> ‘wolves and horses’ and \*w<sup>l</sup>k<sup>w</sup>ōs-k<sup>we</sup> h<sub>2</sub>ekwōs-k<sup>we</sup> ‘both wolves and horses’), cf. Gr. (Hom.) πατήρ ἀνδρῶν τε θεῶν τε ‘father of (both) men and gods’, Ved. *divás ca gmás ca* ‘(both) of heaven and of earth’, *náktam vā hí divā vā vārsati* ‘for it rains (either) by night or by day’. These conjunctions connected phrases, verbs, and sentences as well, cf. Ved. *ā devébhīr yāhi yáksi ca* ‘come with the gods and sacrifice!’ In Rg-Veda (rarely in Greek as well), when connecting what should be vocatives only the first form is actually voc. (the other one is nom.), cf. Ved. *Vāyav Indraś ca. . . ā yātam* ‘O Vayu [voc.] and Indra [nom.], come!’ This reminds one of the pres./inj. (p. 97) and instr. sg./nom. du. ellipsis (p. 63). As in later languages, some pronominal forms (like \*yod ‘that’, \*k<sup>w</sup>od ‘which’, \*k<sup>w</sup>id ‘what’ – p. 88–89) probably functioned as conjunctions already in PIE, cf. Ved. *yád* ‘because; when’, Lat. *quod* ‘(in) that; because; though’, Hitt. *kuit* ‘because’, etc.

## PARTICLES

Particles are usually short and unaccented. Several particles can be reconstructed for PIE, the best known of which is the negative particle \*nē ‘no(t)’ (p. 54). Cf. also the compound \*ne-k<sup>we</sup> ‘and not’ (Lat. *neque* ‘not, and not, also not’, Clb. *nekue* ‘nor, neither, and not’, Hitt. *nekku* ‘not?’, Alb. *nuk* ‘not’). A prohibitive particle \*meh<sub>1</sub> (Ved. *mā*, Gr. μή, Arm.

*mi*, Toch. A/B *mā*, Alb. *mo*, cf. p. 97) can also be reconstructed, originally perhaps the imperative of \**meh*<sub>1</sub>-, cf. Hitt. reduplicated *mi-mm(a)*- ‘refuse’ (Kloekhorst 2008). A couple of other particles can be reconstructed, but often the meaning is not too clear – cf. the emphatic particle \**g*<sup>wh</sup>*e/o* (Ved. *ha, gha*, OCS *že, -go* in the gen. sg. m. pronouns like *togo* ‘that’). In some cases, the particles can appear often after certain forms (cf. the verbal particles that had integrated with verbal forms already in PIE – p. 92), like \**g/ǵe* after personal pronouns, cf. Gr. ἔγωγε ‘I’ – dat. sg. ἔμοιγε, Goth. acc. sg. *mi-k* ‘me’ (but also Goth. *au-k* ‘also’), Hitt. acc. sg. *tu-k* ‘you’ (p. 181), etc.

## INTERJECTIONS

As usual, interjections are often onomatopoeic and can have otherwise non-existent or rare phonological and phonotactical characteristics. In historical view, their phonological development can be irregular, and cognates can be accidental, which makes their reconstruction provisory. Many of them are really simple, like \**ā* ‘ah!’ for surprise or pain (Skr. *ā*, Gr. *ᾶ*, Lat. *ā*, Lith. *à*, Goth. *o*), \**ay* for surprise or pity (Skr. *e, ai*, Gr. *αἶ*, Lat. *ai*, etc.), or exclamation (sometimes used with vocatives) \**ō* ‘oh!’ (Gr. *ὦ*, Lat. *ō*, OIr. *á, a*, Eng. *oh*, etc.). In some cases, the usual sound laws clearly do not apply (for obvious reasons), cf. the laughter onomatopoeia in IE languages (PIE \**ha ha*?): Skr. *ha ha*, Gr. *ἦ ἦ*, Lat. *hahae*, Slav. *ha ha*, Eng. *ha ha*, etc. In others, it (partially) does – cf. Lat. *ēheu* ‘ah, alas!’ (Latin otherwise has no diphthong *eu*) and Skr. *aho* ‘o, oh, alas!’ (with \**ew* > *o*), where PIE \**ēhew* (or \**eHew*?) is reconstructable. A perhaps less trivial reconstruction, but still with many irregular correspondences, is \**way* ‘alas, woe!’ (Gr. *οὐαί*, Lat. *uae*, GAv. *vaiiōi*, Goth. *wai*, W *gwae*, Arm. *vay*).

**Note:** I would like to thank David Mandić, Thomas Olander, Petra Šoštarić, and my students Matija Mužek and Marul Kuljiš for valuable comments on the first draft of the chapter. Of course, all the mistakes are just mine.

## FURTHER READING

The literature on specific problems of PIE morphology is immense – however, there are no modern one-volume monographs dedicated to the whole of PIE morphology as such (or even to the nominal and verbal parts separately). Most PIE comparative grammars (cf. p. 56–57) deal with both phonology and morphology, though their overview of the morphology is often concise in many segments. Of those more recent ones, Szemerényi 1996 is rather detailed, with very helpful explanations of numerous reflexes of morphological forms in separate IE languages, often providing commentary on different views on problems, as well as an extensive (pre-1990) bibliography. However, his views are in some instances dated, as is to be expected. The reconstructions in Beekes 2011 are sometimes idiosyncratic, though some of his heterodox views are very interesting (his take on nominal inflection, such as it is, is presented in Beekes 1985). Fortson 2010 is not too detailed in the PIE section and, for instance, does not provide separate paradigms for all nominal athematic stems. Meier-Brügger 2003 is very useful for his references to literature but also often not too detailed (e.g., he does not list the paradigms in separate IE languages or the reconstructions of all nominal athematic stems), while sometimes dealing too extensively with less important minutiae. A lot of information on the reconstruction of PIE morphology can be found in handbooks on specific branches and languages, e.g., in Sihler 1995, Ringe 2006, or Olander 2015 (with useful references and short surveys of differing views on specific reconstructions). An overview of reconstructed PIE

nouns and adjectives (with their stems) is available in NIL. The problem of PIE gender is tackled in a recent monograph (Matasović 2004) and a collection of articles (Neri & Schumann 2014). A recent monograph covering the IE dual is Fritz 2011. Schmidt 1978 gives an overview of the material and earlier literature concerning personal pronouns, but his reconstructions are rather implausible (cf. Katz 1998 and Kapović 2006 for more recent takes on PIE personal pronouns). For numerals, Szemerényi 1960 is still useful but obviously dated; the Gvozdanović 1991 collection provides a survey of numerals in all IE branches with a number of eminent scholars taking part. Rau 2009 deals with the decades (and also the Caland System). Monographs concerning various aspects of the PIE verbal system are more numerous than works on the nominal system. Hewson & Bubeník 1997 give a useful overview of the verbal systems in various IE branches. LIV is an established handbook on PIE verbal formation, although it has its perks. Clackson 2007 gives an excellent short, problem-based overview of the current issues (the Greco-Aryan model, present/aorist origin, injunctive, optative, middle, etc.) in reconstructing the PIE verbal system. For different famous approaches to pre-PIE origins of the verbal system, cf. Rix 1986 and Jasanoff 2003. The much older Watkins 1969 is still insightful as well, though very controversial. The following monographs deal with specific parts of the verbal system: the aorist (Harðarson 1993 – root aorist; Drinka 1995 – sigmatic; Cardona 1960 – thematic; Bendahman 1993 – reduplicated); the perfect (di Giovine 1990/1996a/b); the middle (Jasanoff 1978, Stempel 1996); the moods (the already mentioned Rix 1986); verbal reduplication (Niepokuj 1997). Further references are available in the mentioned works and chapters in this volume on separate IE branches.

## NOTES

- 1 Hitt. *-uš* may indeed point to \*-ms, not \*-ns (cf., e.g., Kloekhorst 2008: 929).
- 2 With further possible \*-m- cognates in Tocharian (p. 456).
- 3 One must be aware that not all the endings of the words usually taken as examples are always attested. Some actual forms usually cited in full paradigms in handbooks (and sometimes also here) are also not always actually attested as such in the texts (but are assumed according to other attested forms).
- 4 Cf. Matasović 2004: 173–176.
- 5 The original distribution of the thematic \*-ōm (< \*-o-ōm, \*-eh<sub>2</sub>-ōm) and athematic \*-ōm is possibly preserved in (part of) Slavic, cf. the Neo-Štokavian (BCMS) gen. pl. (*o-* and *ā-*stems) *-ā* < PSlav. \*-ǣ < PIE \*-ōm/-eh<sub>2</sub>ōm but gen. pl. (*i-*stems) *-ī* (never \*\*-(i)jā) < \*-ijъ < PSlav. \*-ijъ < PIE \*-ey-ōm (e.g., the *o-*stem gen. pl. *zúbā* ‘teeth’ < PIE \*ǵomb<sup>h</sup>ōm but the old *i-*stem gen. pl. *c’vī* ‘worms’ < PIE \*k<sup>w</sup>ǵmeyōm).
- 6 This did not occur in the gen. pl. ending \*-ōm, because this was originally \*-om in most cases (p. 70). The \*-m did not drop in the Stang-form \*ǵ<sup>w</sup>ōm (p. 71) either.
- 7 The PIE reconstruction here mostly follows the one in Kapović 2006.
- 8 English terminology is rather unfortunate because the traditionally reconstructed PIE categories perfect and imperfect may be confused with aspectually redefined categories of perfective (for traditional “aorist”) and imperfective (for traditional “present”).
- 9 The claim that the Lithuanian original acute accent in the 1 sg. *-ù* (reflexive *-úo-si*) points to a laryngeal is not completely compelling, since Balto-Slavic exhibits the acute accent on some (non-contractional) long-grade endings as well (e.g., in *i-*stem loc. sg. \*-ēy, cf. the Latvian reflexive infinitives in *-tiē-s* from PIE abstract nouns with the loc. sg. \*-t-ēy).



- 10 Cf. Pinault 1992: 153–154.  
 11 It is possible to connect the \*-ey-(e)- (zero \*-i-) suffix with \*-y-e- originally.  
 12 Such injunctives could have been the basis for the making of the Indo-Iranian/Greek imperfect and the Old Irish absolute *berid* ‘carries’ < \*b<sup>h</sup>ereti, *berait* < \*b<sup>h</sup>eronti but conjunct *-beir* ‘says’ < \*b<sup>h</sup>eret, *-berat* < \*b<sup>h</sup>eront, though this is highly controversial (p. 358, 372–373).  
 13 Cf., e.g., Oettinger 1976, Rix 1986, 1988, Kümmel 1996 for the reconstruction of stative (besides the middle).  
 14 Cf. Hermann 1895.

## REFERENCES

- Anttila, Raimo 1969, *Proto-Indo-European Schwebelaut*, University of California Press, Berkeley–Los Angeles
- Beekes, Robert S.P. 1985, *The Origins of the Indo-European Nominal Inflection*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck
- Beekes, Robert S.P. 1988, “The origin of the Indo-European pronominal inflection,” in: Jazayery, Mohammad A. & Winter, Werner (ed.), *Languages and Cultures: Studies in honor of Edgar C. Polomé*, Mouton de Gruyter, Berlin: 73–87
- Beekes, Robert S.P. 2011 [Revised and corrected by Michiel de Vaan], *Comparative Indo-European Linguistics*, John Benjamins Publishing Company, Amsterdam–Philadelphia
- Bendahman, Jadwiga 1993, *Der reduplizierte Aorist in den indogermanischen Sprachen*, Hänsel-Hohenhausen, Egelsbach–Cologne–New York
- Cardona, George 1960, *The Indo-European Thematic Aorists*, Yale University, New Heaven [PhD dissertation]
- Clackson, James 2007, *Indo-European Linguistics. An Introduction*, Cambridge University Press, Cambridge–New York
- di Giovine, Paolo 1990, *Studio sul perfetto indoeuropeo. Parte I: La funzione originaria del perfetto studiata nella documentazione delle lingue storiche*, Dipartimento di Studi Glottoantropologici dell’Università di Roma “La Sapienza”–La libreria Herder, Roma
- di Giovine, Paolo 1996a, *Studio sul perfetto indoeuropeo. Parte II: La posizione del perfetto all’interno del sistema verbale indoeuropeo*, Editrice “Il Calamo”, Roma
- di Giovine, Paolo 1996b, *Studio sul perfetto indoeuropeo. Parte III: Indici*, Editrice “Il Calamo”, Roma
- Drinka, Bridget 1995, *The Sigmatic Aorist in Indo-European: Evidence for the Space-Time Hypothesis*, Institute for the Study of Man, Washington, DC
- Fortson, Benjamin W. IV 2010, *Indo-European Language and Culture. An Introduction*, Wiley-Blackwell, Malden–Oxford
- Fritz, Matthias 2011, *Der Dual im Indogermanischen. Genealogischer und typologischer Vergleich einer grammatischen Kategorie im Wandel*, Winter, Heidelberg
- Gvozdanović, Jadranka (ed.) 1991, *Indo-European Numerals*, Mouton de Gruyter, Berlin–New York
- Harðarson, Jón Axel 1993, *Studien zum urindogermanischen Wurzelaoorist*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck
- Hermann, Eduard 1895, “Gab es im Indogermanischen Nebensätze?,” *Zeitschrift für vergleichende Sprachforschung* 13: 481–535
- Hewson, John & Bubenik, Vít 1997, *Tense and Aspect in Indo-European Languages: Theory, Typology, Diachrony*, Benjamins, Amsterdam

- Jasanoff, Jay 1978, *Stative and Middle in Indo-European*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck
- Jasanoff, Jay 2003, *Hittite and the Indo-European Verb*, Oxford University Press, Oxford
- Kapović, Mate 2006, *Reconstruction of Personal Pronouns in Balto-Slavic with Emphasis on Accentuation*, University of Zadar, Zadar [PhD dissertation]
- Katz, Joshua 1998, *Topics in Indo-European Personal Pronouns*, Harvard [PhD dissertation]
- Kloekhorst, Alwin 2008, *Etymological Dictionary of the Hittite Inherited Lexicon*, Brill, Leiden–Boston
- Kümmel, Martin 1996, *Stativ und Passivaorist im Indoiranischen*, Vandenhoeck & Ruprecht, Göttingen
- LIV = *Lexikon der Indogermanischen Verben*, Rix, Helmut (ed.), Dr. Ludwig Reichert Verlag, Wiesbaden 2001<sup>2</sup>
- Matasović, Ranko 2004, *Gender in Indo-European*, Universitätsverlag Winter, Heidelberg
- Meier-Brügger, Michael 2003 [With contributions by Matthias Fritz and Manfred Mayrhofer], *Indo-European Linguistics*, Walter de Gruyter, Berlin–New York [translation of the 2002<sup>8</sup> German edition]
- Neri, Sergio & Schumann, Roland (eds.) 2014, *Studies on the Collective and Feminine in Indo-European from a Diachronic and Typological Perspective*, Brill, Leiden–Boston
- Niepokuj, Mary 1997, *The Development of Verbal Reduplication in Indo-European*, Institute for the Study of Man, Washington, DC
- NIL 2008, *Nomina im Indogermanischen Lexikon*, Dagmar S. Wodtke, Britta Irslinger & Carolin Schneider, Universitätsverlag Winter, Heidelberg
- Nurse, Derek 2003, “Aspect and tense in Bantu languages,” in: Nurse, Derek & Philippson, Gérard (eds.), *The Bantu Languages*, Routledge, London–New York: 90–102
- Oettinger, Norbert 1976, “Die indogermanische stativ,” *Münchener Studien zur Sprachwissenschaft* 34: 109–149
- Olander, Thomas 2015, *Proto-Slavic Inflectional Morphology. A Comparative Handbook*, Brill, Leiden–Boston
- Pinault, Georges-Jean 1992, “Introduction au tokharien,” *Lalies* 11: 5–224
- Rau, Jeremy 2009, *Indo-European Nominal Morphology: The Decads and the Caland System*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck
- Ringe, Don 2006, *From Proto-Indo-European to Proto-Germanic*, Oxford University Press, Oxford–New York
- Rix, Helmut 1986, *Zur Entstehung des urindogermanischen Modusystems*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck
- Rix, Helmut 1988, “The PIE middle: Content, forms and origins,” *Münchener Studien zur Sprachwissenschaft* 49: 101–119
- Schmalstieg, William R. 1998, “The instrumental singular origin of the nominative dual and plural,” *Linguistica Baltica* 7: 179–186
- Schmidt, Gernot 1978, *Stammbildung und Flexion der indogermanischen Personalpronomina*, Harassowitz, Wiesbaden
- Sihler, Andrew 1995, *New Comparative Grammar of Greek and Latin*, Oxford University Press, New York–Oxford

- Stempel, Reinhard 1996, *Die Diathese im Indogermanischen: Formen und Funktionen des Mediums und ihre sprachhistorischen Grundlagen*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck
- Szemerényi, Oswald J.L. 1960, *Studies in the Indo-European System of Numerals*, Carl Winters Universitätsverlag, Heidelberg
- Szemerényi, Oswald J.L. 1996, *Introduction to Indo-European Linguistics*, Clarendon Press–Oxford, Oxford–New York [translation of the 1990<sup>4</sup> German edition]
- Watkins, Calvert 1969, *Indogermanische Grammatik*. III/1. Geschichte der Indogermanischen Verballexion, Carl Winter Verlag, Heidelberg

# PROTO-INDO-EUROPEAN SYNTAX

*Thomas Krisch*

After general remarks mentioning and evaluating older literature and the data basis for Proto-Indo-European (henceforth PIE) syntax, this article will describe major fields of present day research. Clause alignment and the ergative hypothesis are not dealt with in this article (see the treatment by Matasović, p. 162–164). The end of this article contains a short conclusion.<sup>1</sup>

## GENERAL REMARKS

### **A very short survey of older literature on PIE syntax**

Berthold Delbrück (1842–1922) is the “father” of comparative PIE syntax. His monumental work on comparative syntax in three volumes encompasses phenomena that we now subsume under morphosyntax and syntax. Among them are the use of nominal number, case (including matters of valency of the verb), prepositions, the alternations between adjectives and nominal genitives, the use of tenses, moods and genera verbi and the organization of sentence structure (word order, elliptical constructions, appositions, agreement issues, question sentences, relative clauses). Delbrück analyses numerous examples from ancient Indo-European languages (AIELs), and he states that if constructions and meanings coincide in the AIELs, this could mean that the same was true also for PIE (Delbrück 1967=1893: 86). This statement still makes sense today.

The syntactic framework of Delbrück and other eminent researchers like Jacob Wackernagel (e.g. Wackernagel (1920 and 1924), outstandingly translated and edited by Langslow (2009)) was based on traditional Latin grammar and philological interpretation. The next methodical landmark was an article in the spirit of structuralism by Calvert Watkins (1963) about word order in Old Irish with archaic verb-final position (which Watkins expects to reflect the PIE state). These are cases of Bergin’s Law (Watkins 1963: 24) and of tmesis. In Irish tmesis, the preverb, optionally followed by Wackernagel clitics, introduces the sentence, and the finite verb ends it. Assuming a univerbation process where the verb moves from the final position to “its” preverb in sentence-initial position, Watkins succeeds in integrating the deviating (unmarked verb-initial order) Old Irish data into a general picture of PIE with verb-final as the unmarked order. The preverb could not move to the end of the sentence because that would violate Wackernagel’s Law in case clitics are attached to it. This reasoning is still convincing and very modern. The first part of McCone’s outstanding dissertation (1979) takes ideas from Watkins 1963 and describes sentence patterns in Hittite, Vedic and Greek with Watkins’ formal devices.

The 1970s laid the ground for much of what is still discussed today in studies on PIE syntax. They brought typology (in the Greenberg 1966 tradition, cf. also Matasović,

p. 165) and theory to bear on the description of PIE syntax, with an emphasis on word order: Lehmann (1974) reconstructs \*SOV, Friedrich (1975) thinks \*SVO and \*SOV are equally important, and Miller (1975) considers \*VSO as well. Referring to this typological approach, Watkins (1976: 305) deploras “. . . that the rebirth of Indo-European syntax has taken place in the bed of Procrustes”. This criticism is true as long as one looks at the word order typology as rigid order (as was partly done in the cited literature) and wants to sweep the counterexamples under the carpet. The development of research in the last 40 years has brought some progress in this respect; see our section on word order. Another criticism of Watkins, however, is still to be kept in mind today: He suspects that the picture of PIE syntax could also be influenced by “Teeter’s Law” (the language of the family you know best always turns out to be the most archaic; cf. Watkins 1976: 310). I think that every researcher (including myself, of course) faces this problem, but this does not have to be evaluated in a purely negative way. Also, the chance of getting precious new insights is thus increased.

A very influential article of the 1970s was Dressler 1971. In that paper, the author mentions a number of important fields of research on PIE syntax, which are, among others, still elaborated in current research: typological issues (cf. also Dressler 1968), the syntax of case (see our sections on case functions and on argument structure), sentence prosody (see our section on Wachernagel’s law), word order and text syntax (see our section on verb positions with arguments taken from text syntax).

Today’s research on PIE syntax encompasses a broad spectrum with regard to contents. Most of it shows a sharpened awareness of methodology and search for points of contact with general linguistic ideas. This new perspective is an outcome of research done in the 1960s and 1970s.

## The data

The comparative method of reconstructing the PIE language system has been very successful in the realm of phonology, morphophonology (especially ablaut) and morphology. It was developed in the 19th century (e.g. Franz Bopp (1791–1867), Jacob Grimm (1785–1863) and the Neogrammarians like Karl Brugmann (1849–1919)) and is still being refined on today in numerous publications. Using this method, one collects probable cognates (units in form and function, normally words) in a number of related languages, investigates regular correspondences between them and finally reconstructs proto-forms. These, in turn, give rise to sound correspondences, to morphophonological theories of nominal and verbal inflection and derivation (accent-ablaut types), to the reconstruction of compound types, etc.

It is very difficult to apply this comparative method to syntax since the oldest texts in a number of ancient Indo-European (IE) languages do not even share a type of text. In Hittite, one finds prose texts with laws, narratives and rituals. In Greek (apart from the Mycenaean inventories), one has the great monument of Homer, a poetic text; prose texts are attested only in later times. In Vedic and Avestan, we have metrical texts with ritualistic content, etc., etc. There are no primary data for PIE sentences available, and there are no attestations of exactly the same sentence with all words cognate in more than one language in the AIELs with the exception of frozen syntagms (see below) in the language of poetry. The attempted material reconstruction of “real” PIE sentences normally reveals only the state of research for the respective reconstruction, and therefore one cannot use it as primary data, of course.<sup>2</sup>

In the case of attested texts in ancient languages, one at least has access to some data which have been handed down and have survived time largely by chance. Even these

data suffer from big disadvantages when compared with linguistic data from living languages. One cannot freely form sentences which can be judged by native speakers as to whether they are grammatical or ungrammatical. Normally syntactic reconstruction can just reconstruct schemata, abstract sentence structures, and thus differs from phonological, morphological and lexical reconstructions with more “substance”. But there are very rare cases where one can reconstruct, e.g., word order including “substance”, with the same (ideally etymologically related) words occurring in more than one language. This could offer a “punctum Archimedis”, an Archimedean starting point. Let us start with the one which to my knowledge is the best syntagm that can be reconstructed with “substance” for PIE, namely that a hero is killing a dragon (cf. also Watkins 1995, esp. chapter 27 and 28; Krisch 2001). There is good evidence to reconstruct this syntagm as offered in (1), with government of the verb to the right (data from Vedic, Avestan and Greek (see the examples in (2a–d)):

- (1) PIE  $*(\acute{e})-g^{wh}en-t$   $*(H)e/og^{wh}-im$   
 (Augment)-kill-3SG.INJ.PRES./ (IMPF.) serpent-ACC.SG.
- 
- ‘He kills / (killed) the serpent’
- 
- (2) a) Vedic: RV 1,32,2 (*áhann* < PIE  $*e-g^{wh}en-t$ ; *áhim* < PIE  $*(H)e/og^{wh}-im$ )  
*áhann* *áhim*  
 kill.3SG.IMPF. serpent.ACC.SG.
- 
- ‘he killed the serpent’
- 
- b) Avestan: Y.9,8 (*janat* < transposed PIE  $*g^{wh}en-e-t$ ; *ažim* < PIE  $*(H)e/og^{wh}-im$ )  
*yo* *janat* *ažim* *dahakam*  
 who kill.3SG.INJ.PRES. serpent.ACC.SG. Dahāka.ACC.SG.
- 
- ‘who killed Aži (“the dragon”) Dahāka’
- 
- c) Greek: Pindar P. 10,46 (*ἔπεφνε(ν)* < PIE  $*e-g^{wh}e-g^{wh}n-e$  reduplicated aorist; Gorgon: lit. ‘the terrible one’; winged daemon with serpentine locks in the hair)  
*ἔπεφνέν* *τε* *Γοργόνα*  
 kill.3SG.AOR. and Gorgon.ACC.SG.
- 
- 
- d) Pindar, P. 4,249: [*κτεῖνε* < PIE  $*tkenye^{\circ}/kpenye^{\circ}$  (injunctive present) or < PIE  $*tken-s^{\circ}/kpen-s^{\circ}$  (injunctive aorist); *ὄφιν* < PIE  $*(H)e/og^{wh}-im$ ]  
*κτεῖνε* . . . *ὄφιν*  
 kill.3SG.IMPF. / AOR. serpent.ACC.SG.
- 
- ‘he killed the serpent / dragon’

If one combines the Greek sentences (2c) and (2d), one gets Greek *ἔπεφνε(ν) ὄφιν*, which can be projected back to  $*e-g^{wh}e-g^{wh}n-e *He/og^{wh}im$ . Except for the tense of the verbal form, this reconstruction is the same as the one in (1), which was arrived at by the Vedic and Avestan forms.

But from this evidence one must not draw the conclusion that the ancient languages in question and PIE were necessarily languages with a VO order. This is confirmed by the fact that the inverse order (OV) of the same syntagm is attested in a number of ancient IE languages as well and can be reconstructed for the proto-language preceding Vedic, Greek and Hittite (which of course is PIE as well). Cf. the reconstruction in (3) with government to the left:

- (3) PIE \*(H)e/og<sup>wh</sup>-im                      \*(é)- g<sup>wh</sup>en-t  
 serpent-ACC.SG.                      (Augment)-kill-3SG.INJ.PRES./((IMPF.)  
 ←  
 ‘He kills /(killed) the serpent’

This reconstruction is supported by syntagms attested in Hittite, Vedic and Greek (see the examples in (4a–d)):

- (4)  
 a) Hittite KUB 17.5 Vs I (Middle Hittite): (*kuenta* < PIE \*g<sup>wh</sup>en-t(o))  
 (17) <sup>DIM</sup>aš ú-it nu-kán <sup>MUŠ</sup>il-lu-[ya-an-ka-an] (18) ku-en-ta DINGIR<sup>MES</sup>-ša  
*kat-ti-iš-ši e-še-ir*  
<sup>DIM</sup>aš              úit                      nu=kán              <sup>MUŠ</sup>illu[yankan]              (18) *kuenta*  
 Weathergod come.3SG.PRET. SC=LOC.PTCL Illuyanka.ACC.SG.              kill.3SG.PRET.  
 ←  
 DINGIR<sup>MES</sup>-ša = *a*                      *katti=šši*                      *ešer*  
 God.NOM.PL. = but                      with=him (lit. POSS.PRON.)              be.3PL.PRET.  
 ‘The weathergod (i.e. Tarhunnaš) came and killed (the serpent) Illuyanka; but the gods were with him’

- b) Vedic RV 2,15,1 (*áhim* < PIE \*(H)e/og<sup>wh</sup>-im; *jaghāna* < PIE \*g<sup>wh</sup>e-g<sup>wh</sup>on-e)  
*trikadrúkeṣu*              *apibat*              *sutasya*              *asyá*              *máde*  
 Trikadrúka-day.LOC.PL. drink.3SG.IMPF. Soma.GEN.SG. he.GEN.SG. intoxication.LOC.SG.  
*áhim*                      *indra*                      *jaghāna*:  
 serpent.ACC.SG.              Indra.NOM.SG.              slay.3SG.PERF.  
 ←  
 ‘At the Trikadrúka-days he drank from Soma; in the intoxication (caused) by this (one), Indra has slain the serpent’

- c) Greek: H. Ap. 300-301 [κτείνε < PIE \*t<sup>h</sup>ken-ye°/k<sup>h</sup>pen-ye° (injunctive present) or < PIE \*t<sup>h</sup>ken-s°/k<sup>h</sup>pen-s° (injunctive aorist)]  
 ἐνθα              δράκαιναν /              κτείνεν              ἄναξ              Διὸς              υἱὸς  
 there              she-dragon.ACC.SG. kill.3SG.IMPF./AOR. lord.NOM.SG. Zeus.GEN.SG. son.NOM.SG.  
 ←  
 ἀπὸ                      κρατεροῖο              βιοῖο  
 with (lit. from)              strong              bow.GEN.SG.  
 ‘there, the lord (i.e. Apollo, TK), the son of Zeus, killed the (female) dragon with the strong bow’

- d) Greek: Pindar, O. 13,88 ff. (ἐπεφνε(v) < PIE \*e-g<sup>wh</sup>e-g<sup>wh</sup>n-e reduplicated aorist) καὶ Χίμαιραν πῦρ πνέοισαν καὶ Σολύμους ἔπεφνεν  
 ←  
 ←

καὶ	Χίμαιραν	πῦρ	πνέοισαν	καὶ
and	Chimaira.ACC.SG.F.	fire.ACC.SG.F.	breathe.PTCP.AOR.ACC.SG.F.	and
Σολύμους	ἔπεφνεν			
Solymoi.ACC.PL.	kill.3SG.AOR.			

‘And he (i.e. Bellerophontes, TK) killed the fire-breathing Chimaira and the Solymoi’

The Hittite example in (4a) shows a cognate verb form that fits with the reconstruction in (3), and the Vedic instance in (4b) is (with the exception of tense) almost an exact equivalent to (3). In the Greek example (4c), the ‘dragon’ word and the verb have only a semantic correspondence to the reconstruction in (3), and in (4d) only the verb is etymologically related to the reconstruction, and the object shows words and names for dragons with etymologies different from \*(H)e/og<sup>whi</sup>-. But, nevertheless, a word for ‘dragon’ is put in the accusative case in both Greek examples. This is as close as one can get when one wants to reconstruct genuine PIE sentences.

The reconstructions in (1) and (3) seem to contradict each other. A number of conclusions from this fact are possible, e.g. the following:

- IE and the oldest languages were fluctuating between OV and VO. This would describe the fact that our sentence appears in both varieties in Greek and in Vedic Sanskrit. Note, however, that Hittite shows no variant with VO. This scenario could lead to a project which tries to disentangle the OV and VO features of the ancient IE languages and the reconstructed proto-language, using, e.g., typological criteria like those of Greenberg (1966) and his followers.
- IE was like Hittite, which, as Luraghi has shown in her seminal book from 1990, is a fairly strict OV language. Note also that our sentence in Hittite shows no variant with VO. The languages with VO order for this sentence, namely Greek, Sanskrit and Avestan, could reflect a younger state, a shift to VO. If one adopts this scenario, this could result in looking for layers of OV or VO in the oldest history of the languages showing variation and depicting developments from OV to VO in these languages.
- The IE base-type was OV, but under certain pragmatic conditions it allowed for a movement of the verb to the front of the sentence. This is the view I will adopt.

We have seen that frozen syntagms in cognate languages can lead to revealing conclusions for PIE syntax and may lead to further investigation. Most of these frozen syntagms go back to poetic language. An up-to-date overview of all aspects of IE poetic language including syntax (using a generative framework) is Hajnal 2008. An extensive collection of text correspondences and of frozen syntagms in AIELs can be found in Schmitt 1967.

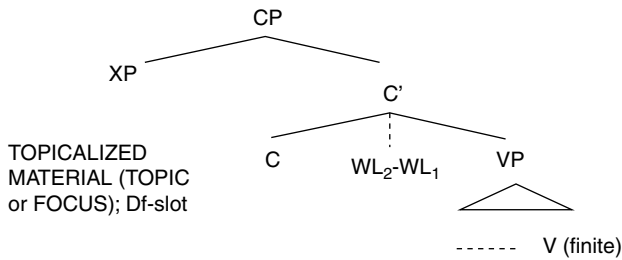
## MAJOR PRESENT FIELDS OF RESEARCH ON PIE SYNTAX

### Word order

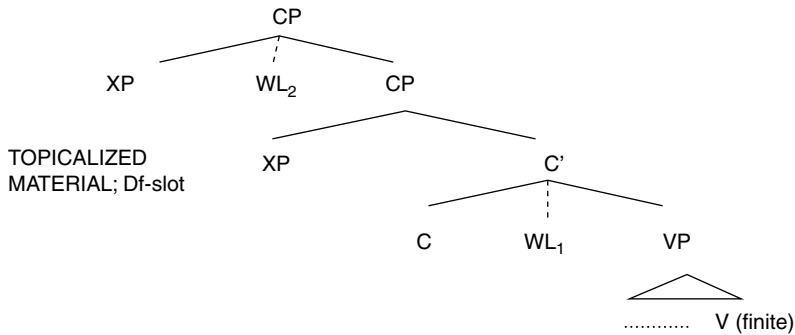
Inspired by Mark Hale (1987) and Paul Kiparsky (1995) and expanding some earlier ideas of Krisch 1990, I have used generative grammar (the Government and Binding model) to describe word order in AIELs and in PIE in a number of publications. The model I use now is slightly different from the models I have used in earlier publications, but it is



(5) cf. also Krisch 2015

**FIGURE 1.2** SENTENCE SCHEME 1

(6) cf. also Krisch 2015

**FIGURE 1.3** SENTENCE SCHEME 2

compatible with the older model. Following Keydana (2011), I have integrated the label Df-slot, and I have renumbered the Wackernagel clitics following Keydana (2011) and Lühr & Zeilfelder (2011). For practical purposes (mainly motivated by the treatment of Wackernagel particles) I assume two schemes of phrase structure: (5) and (6). The two schemes are compatible with Government and Binding (the labels form a subgroup of the universal set of phrase structure), and they are in fact only one structure: The second scheme is derived from the first one by “Chomsky adjunction”. For the abbreviations used see the appendix above the bibliography. The Df-slot (discourse functional slot) is a “landing-site” for topical and focal material. The number of categories used is held as small as possible. The morphology is not done in a generative way; the word forms are inserted fully inflected into the syntactic tree.

One can use these schemes for a number of phenomena, which I will shortly characterize in the following sections.

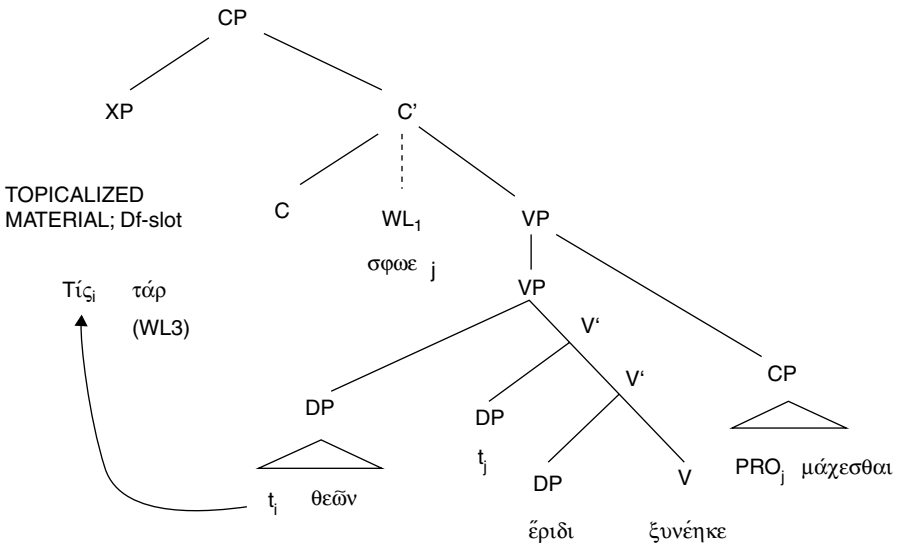
#### *Discontinuous constituents (hyperbaton/tmesis)*

One of the most striking features of AIELs are discontinuous constituents. Most of them (except for movements inside DP and PP) can be quite simply described by moving parts

of constituents to the left or the right of the VP. In (7) (Scheme 1), the interrogative pronoun τίς is fronted out of its DP to the front, into the Df-slot, which is not surprising for a clear focal element. In this case, only Wackernagel particles intervene (cf. also Krisch 1998: 372).

- (7) Hom. Il. 1,8:  
 Τίς τάρ σφωε θεῶν ἔριδι  
 Who.NOM.SG. there they.ACC.DU. god.GEN.PL. quarrel.DAT.SG.  
 ξυνέηκε μάχεσθαι  
 bring.together.3SG.AOR. fight.INF.MID.  
 “Who on earth<sup>3</sup> of the gods brought both of them together to fight in quarrel?”

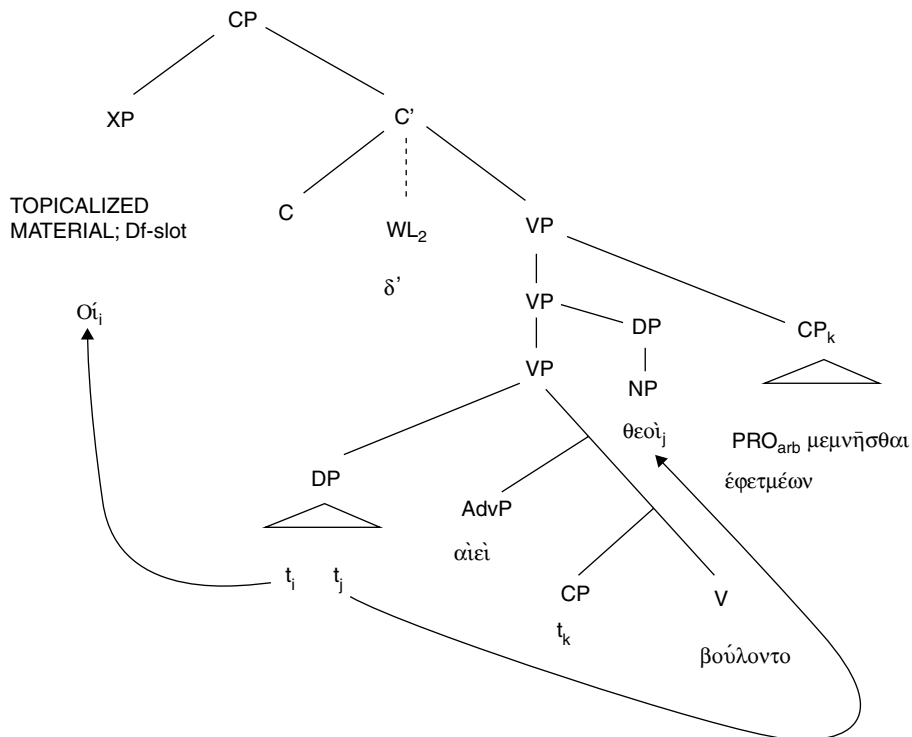
FIGURE 1.4 STRUCTURE OF HOMER, ILIAD 1,8



In (8), part of the DP is fronted, and part of it is extraposed to the right of the VP (by adjunction): The Homeric article/determiner οἱ, which has stronger demonstrative value than the article of later Greek (cf., e.g., Chantaine 1981=1953: 158), and which conveys anaphoric-topical information, is put into the Df-slot to the left, whereas the head of the NP, θεοῖ, is extraposed to the right.

- (8) Hom. Od. 4,353:  
 οἱ δ' αἰεὶ βούλοντο θεοῖ  
 The(se)NOM.PL. but always want.3PL.INJ.MID. god.NOM.PL.  
 μεμνήσθαι ἐφετμέων  
 remember.INF.PERF.MID. command.GEN.PL.  
 “But the(se) gods always want to remember their commands”

FIGURE 1.5 STRUCTURE OF HOMER ODYSSEY 4,353



Hyperbata are widespread among AIELs, and thus they are a syntactic device which can easily be reconstructed: In (9) one finds examples for this phenomenon in Vedic (*ūtīm sahasrasātāmāṃ* is extraposed), (10) is an example from Latin (*pulmoneum* is fronted), and there are also examples in Hittite ((11); the genitive is extraposed to the right).

(9) Vedic RV 1,10,10:

*vṛśantamasya* hūmaha *ūtīm*  
 extremely.bull-like.one.GEN.SG. call.1PL.PRES.MID. help.ACC.SG.F.  
*sahasrasātāmāṃ*  
 most.winning.a.thousand.ACC.SG.F.

‘We ask for the most bull-like one’s (i.e. god Indra’s, TK) help, which excels in winning a thousand’

(10) Old Latin Plaut. Rud. 511:

*Pulmoneum* *edepol nimis*  
 Consisting.of.lung.ACC.SG.M. by.Pollux! too much  
 uelim *uomitum* uomas  
 wish.1SG.SUBJ. vomit.ACC.SG.M. vomit.2SG.SUBJ.

‘I wish very much that you vomit a vomit consisting of lung particles.’

- (11) Hittite StBoT 8, Vs I 22 (Old Hittite ritual) (Otten & Souček 1969: 20)

II        <sup>D</sup>**Hantasepuš**                      *ḫarwani*                      **GI[(Š-aš)]**  
 Two Hantasepa-deities.ACC.PL.    have.1PL.PRES.    wood.GEN.SG.  
 ‘We have two Hantasepa-deities (made) of wood.’

Normally, the phenomenon of hyperbaton/tmesis is looked upon as a poetic device. But it is also a feature which appears in everyday speech. This is shown by the Plautus example in (10) and by the fact that inscriptions also show the phenomenon (cf. (12), where the location from which Dropides comes (the adjective Αφιδνατος) is moved to the right)

- (12) Greek IG II III
- <sup>2</sup>
- 2636 (Attika) (cited after Pfohl 1980: 169):

[**Δ**]ρ**οπίδης**                      κατέλαβε                      [**ἀ**]νασάξιμον  
 Drōpídēs.NOM.SG.M.    take.possession.of.3SG.AOR.    equipped.again.ACC.SG.  
 μέταλλον    **Αφιδνατος**  
 mine.ACC.SG.    from.Ap<sup>h</sup>idna.NOM.SG.M.  
 ‘Dropides from Ap<sup>h</sup>idna took possession of a mine which was restarted’

A surely old PIE construction pattern is a sentence where a preverb (originally adverb) is fronted, thus appearing apart from “its” verb. This is usually called “tmesis” (Gr. τμήσις lit. ‘cutting’). Already Wackernagel observed this for Greek: “from the earliest period, tmesis . . . is commonest when the preverb is in clause-initial position” (Wackernagel 1924: 174 = Langslow 2009: 616). Actually, in this case, for PIE the term “tmesis” is misleading. The most probable scenario for the oldest texts and for PIE is that the “preverb” had only a loose connection to the verb and has to be looked upon more as a (normally local) adverb (cf. also Krisch 1984, especially chapter 3), which can appear in sentence-initial position (Df-slot) to give a frame to the sentence. Of course, the sentence-initial position is not the only position of these adverbs in PIE and in the AIELs, but as already sketched above, Watkins (1963) has taught us that this construction helps to understand the development of the Celtic verb-initial construction. Examples (13)–(15) offer some examples for this construction in Hittite, Vedic and Greek. In all three examples, the preverb can be understood as an independent adverb or as a preverb. For the Hittite example in (13) note that, e.g., German has a prefixed verb *hinterherrufen* ‘to call after someone’ with the prefix *hinterher* ‘behind’.

- (13) Hittite StBot 8 (Old Hittite ritual) Rs. III,4 (Otten & Souček 1969: 30–31):

**appananda**=*ma=šše*                      *kē*                      *mēmahḫi*  
 behind.after=but=he.DAT.SG.    this.ACC.SG./PL.    say.1SG.PRES.

‘but from behind I say this to him (i.e. to the eagle who is set free, TK) / but after him (i.e. the eagle set free, TK) I say this’

- (14) Vedic RV 6,53,4

*ví*    *pathó*                      *vājasātaye*                      *cinuhí*  
 away path.ACC.PL.    obtaining.good.DAT.SG.    arrange.2SG.IMP.PRES.

*ví*    *mṛdho*                      *jahi*  
 away adversary.ACC.PL.    strike.2SG.IMP.PRES.

‘open up paths for obtaining goods, slay enemies’

(15) Greek Hom. Il. 1,308–11:

ἐν δ'	ἑρέτας	ἔκρινεν	εἴκοσιν, ἐς δ'	ἑκατόμβην /
In but/and	rower.ACC.PL.	chose.3SG.IMPF.	twenty in but/ and	Hekatombs.ACC.SG.

βῆσε	θεῶ	ἄνῃ	δὲ	Χρυσήϊδα /
make.to.go.3SG.INJ.AOR.	god.DAT.SG.	onto	but/and	Chryses.ACC.SG.F.
καλλιπάρηον	εἶσεν	ἄγων		
fair.cheeked.ACC.SG.M./F.	make.sit.3SG.AOR.	lead.PTCP.PRES.NOM.SG.M.		

‘(But the son of Atreus launched a swift ship on the sea,) and chose for it twenty rowers, and drove on board a hecatomb for the god, and brought the fair-cheeked daughter of Chryses onto it and made her sit.’

### Wackernagel’s Law

In 1892 Wackernagel published his long famous article “Über ein Gesetz der indogermanischen Wortstellung”. He states (elaborating earlier work by Delbrück and Bartholomae) that certain clitics tend to occupy the second position in the PIE sentence. He mainly bases this claim on data taken from Ancient Greek, but he also deals with a small amount of Latin, Vedic and Old Iranian and Celtic material. Already eight years later, Delbrück integrated the law into his handbook (Delbrück 1967=1900: 49–50). The following remark by Watkins dating from 1964 is still valid today: “. . . one of the few generally accepted syntactic statements about IE is Wackernagel’s Law, that enclitics originally occupied the second position in the sentence” (Watkins 1964: 1036). Decades after Wackernagel, this law has also been confirmed by newly discovered AIELs, especially the Anatolian languages (see the Hittite examples in Krisch 1990) and Mycenaean, thus making it a robust candidate for reconstruction. In the following examples, the clitics in Wackernagel position are in boldface. For reasons why some of the Greek clitics (see (16)) carry an accent and nevertheless “count” as clitics, see Krisch (1990: 75–76).

(16) Greek Hom. Il. 18,394:

ἦ	<b>ῥά νύ</b>	<b>μοι</b>	δεινῆ	τε	καὶ	αἰδοίῃ
in.truth	then now	I.DAT.SG.	terrible.NOM.SG.F.	and	also	honorable.NOM.SG.F.
θεῶς		ἔνδον				
goddess.NOM.SG.		inside				

‘then, **in truth, now** the terrible and honorable goddess (is) inside **for me**’

(17) Vedic RV 1,175, 2

ā	<b>nas</b>	<b>te</b>	gantu	matsaró
here	we.ACC.	you.GEN.SG.	come.3SG.IMP.AOR.	bull-like.NOM.SG.M.
vjśā			mádo	váreṇyah
intoxicating.NOM.SG.M.			drunkenness.NOM.SG.M.	excellent.NOM.SG.M.

‘The intoxicating, bull-like, excellent drunkenness **of you** (i.e. God Indra, TK) shall come **to us**.’

(18) Hittite (Held 1957: 14) KUB XIII 2 III 16:

*kuiš=*            *an=*            *šan*            *EGIR-pa*    *tarnai*            *n=*            *an*  
 who.NOM.SG.    he.ACC.SG.    LOC.PTCL.    back            let.3SG.PRES.    SCONN.    he.ACC.  
*šakuwanzi*  
 lock.up.3PL.PRES.  
 ‘(Anyone) who lets him (come) back they (shall) put in prison’

Wackernagel’s Law has become clearer through the research of the last few decades. This research is mainly based on Vedic, which probably can be taken as a model for PIE. Now one divides the clitics into three groups (cf. also Keydana 2011). Only the first two are genuine Wackernagel clitics.

- WL<sub>1</sub> clitics (clitic pronouns) typically occupy the second position in the sentence, but in sentences with wh-words they always follow the wh-word (see examples (21) and (22))
- WL<sub>2</sub> clitics (sentence connectors and sentence adverbs) always occupy the position after the first word of the sentence
- WL<sub>3</sub> clitics are adjacent to the word or constituent they take scope over. Typically, these are emphatic particles. If their host is put at the front of the sentence, they also occupy the Wackernagel position (cf. already Hale 1987: 44). Lühr (2009) deals with such particles from an information structural point of view.

In the Ṛgveda, the following (types of) data have to be described (cf. Keydana 2011: 122):

(19) RV 1,176,4

*(ásunvantam samam jahi dūṇāśam yó ná te máyah |) asmábhyam asya védanam  
 daddhí (sūriś cid ohate) ||*  
 ‘(Kill everyone who does not press soma, the inaccessible one, who is not your  
 delight) Give to us his property! (He boasts (being) an inciter (of a sacrifice))’  
*asmábhyam asya védanam daddhí*  
 we.DAT.            this.GEN.SG.    property.ACC.SG.    give.3SG.IMP.PRES.

As the context (19) shows, the accented pronoun *asmábhyam* is in contrastive focus to the *ásunvant-*, the ones not pressing soma. Therefore, it stands in the Df-slot (cf. (5)) and the WL<sub>1</sub> *asya* is immediately following.

(20) RV 1,76,1

*Kéna vā te mánasā dāśema*  
 Which.INSTR.SG.    or    you.DAT.SG.    spirit.INSTR.SG.    sacrifice.1.PL.OPT.PRES.  
 ‘With which spirit should we sacrifice you?’

In (20), we see that the WL<sub>2</sub> clitic *vā* precedes the WL<sub>1</sub> clitic *te* (the wh-word + WL<sub>2</sub> clitic form a prosodic word to which the WL<sub>1</sub> clitic attaches).

(21) RV 6,27,1

*indraḥ kím asya sakh<sub>yé</sub> cakāra*  
 Indra.NOM.SG.    what.ACC.SG.N.    this.GEN.SG.M.    friendship.LOC.SG.    do.3SG.PERF.  
 ‘What has Indra done in his (i.e. Soma’s, TK) friendship?’

Example (21) shows a remarkable feature of the WL<sub>1</sub> clitics: they follow the wh-word, even if another word/constituent precedes it. This construction can also be found in Greek (scheme 2):

(22) Od. 20,47–48

(αὐτὰρ ἐγὼ θεός εἰμι), διαμπερὲς ἦ σε φυλάσσω ἐν πάντεσσι πόνοισ’  
 ‘(But I am a goddess,) (I) who incessantly protect you in all toils’

διαμπερὲς ἦ σε φυλάσσω ἐν πάντεσσι πόνοισ’  
 incessantly WHO.NOM.SG.F. you.ACC.SG. protect.1SG.PRES. in all.DAT.PL. toil.DAT.PL.

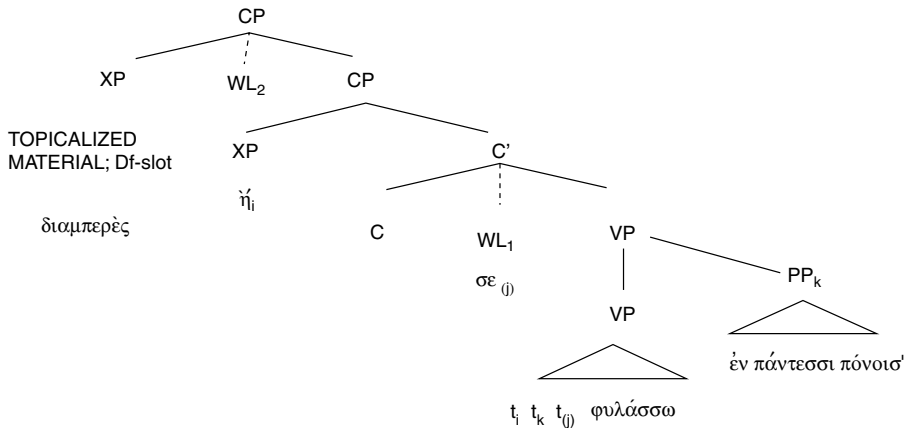


FIGURE 1.6 STRUCTURE OF HOMER, ODYSSEY 20, 48

These facts have led to a number of proposals for how to characterize them. Hale (1987) is not very explicit about whether Wackernagel's Law is a syntactic process (a movement operation leaving a trace) or another type of process interacting with syntax. He describes example (21) and others in the following way: "It is clear from these examples that the WL clitics take second position, defined before the topicalization but after WH-movement places *ká-* (i.e. the wh-question word, TK) in COMP" (Hale 1987: 42). Applied to (21) and to our model in (2) this means that *kim* is in the XP position of the second CP (like *hĕ* in (22)), then WL clitics "take" second position, and finally *indraḥ* is topicalized. "Take" could either mean a syntactic operation or some kind of intervening prosodic operation. Keydana (2011: 115) criticizes this account by calling attention to the fact that an example like (19) cannot be described this way. The WL<sub>1</sub> clitic *asya* could not appear in this position because it cannot move there before topicalization.

In a later paper, Hale abandons rule ordering and describes Wackernagel clitics as "syntactic movement to C<sup>o</sup>" (Hale 1996: 192). I agree with Keydana (2011: 119) that such a movement appears to be strange. The base position of the pronouns should be a full XP, and movement from an XP to a head position (C<sup>o</sup>) is disallowed by the concept of the structure preservation principle (cf. Haegeman 1994: 338). But, on the other hand, there are examples with clitics of the WL<sub>1</sub> type far away from the left periphery (cf. the clitic personal pronoun *t<sub>u</sub>vā* in (23)):

(23) RV 1,45,5 (cf. Hale 1995: 263)

*(ghj'tāhavana sant,yemā u śú śrudhī girah |)*

*yābhiḥ kānvasya sūnāvo hāvante āvase t<sub>v</sub>vā ||*

‘[you, the one to whom the ghee oblation belongs, honest one, listen well to these songs of praise] with which the sons of Kanva invoke you for help’

*yābhiḥ kānvasya sūnāvo hāvante āvase t<sub>v</sub>vā*

which.INSTR.PL. Kanva.GEN.SG. son.NOM.PL. invoke.3PL.PRES. help.DAT.SG. you.ACC.SG.

Hock’s (e.g. 1996) account of the Vedic left periphery offers a prosodic (not syntactic) template solution. The idea of prosody is taken up by Keydana (2011), who analyses the WL<sub>2</sub> clitics as prosodically attached to the first word in the sentence, whereas the WL<sub>1</sub> clitics always stand at the right of the left periphery, having the prosodic prefield of the sentence as their host.

In my model (see above, examples (5) and (6)) these “Wackernagel particles” occupy positions which are marked with dotted lines, which means that I think their syntactic status is rather loose, representing an interface between syntax and prosody. This solution could suit Keydana’s idea of insertion of the clitics into the prosodic prefield of the sentence (in my model: before VP) but also allows taking into account that the WL<sub>1</sub> clitics can occupy an argument position in the sentence. For the last point see coindexed  $\sigma\epsilon$  in parentheses in (22). Cf. also (23), where prosody attaches the clitic *t<sub>v</sub>vā* to tonic *āvase* at the end of the sentence. In (23) the syntactic base position of *t<sub>v</sub>vā* would probably be a position in front of the verb (a fully fledged DP would occupy this position). Thus, also in the case of (23), the prosody and syntax interact. Coindexing  $\sigma\epsilon$  in parentheses in (22) should indicate that I do not assume a movement operation here which is motivated by syntax alone.

A view completely different from the ideas just presented is promoted by the phonologists Agbayani and Golston (2010), who work with the concept of postpositives (Dover 1960) like the clitic \*k<sup>w</sup>e (= Lat. *-que*, Gr. *-τε*) ‘and’ or (in their view) non-clitic Greek *-δέ* that cannot begin a phonological phrase. In their analysis, conjunctions (also the postpositive ones) originate syntactically between the conjoined phrases or clauses, and in the case of a postpositive conjunction like \*k<sup>w</sup>e an adjacent (phonological) word of the second conjunct is moved to the left of it (a phonological, not a syntactic process). In the case of clauses which are not conjoined, but would syntactically start with postpositives, the first (phonological) word of the second conjunct would also be moved to the left. An example of the last case is (24): *κεν* is a postpositive clitic particle denoting a “possible world” (cf. Krisch 1986: 18), and *μιν* is a postpositive clitic personal pronoun. According to Agbayani and Golston, the sentence (without a conjunction) would start syntactically with two postpositives, and phonology would trigger movement of *τότε* ‘then’ to the front of them, since postpositives cannot begin a phonological phrase. But this would mean that (syntactically) the Wackernagel clitics *κεν* (WL<sub>2</sub>) and *μιν* (WL<sub>1</sub>) do not occupy the second position but rather that the sentence starts with *κεν* in the syntactic component.

(24) Hom. II. 1,100 (Agbayani & Golston 2010: 18;  $\phi$  = phonological phrase; the authors underline the word that is moved and indicate the position from which it has moved with underlining as well)

*(τότε κέν μιν \_\_\_ ἱλασσάμενοι)<sub>\phi</sub> (πεπίθουμιν)<sub>\phi</sub>*

then PTCL. him appeasing let.us.persuade

‘then let us persuade him by appeasing him’

But, as far as I see, this interesting approach does not really solve the problem of the different behaviour of WL<sub>1</sub> and WL<sub>2</sub> clitics as outlined above.



*The position(s) of the verb in AIELs and in PIE*

The author of this chapter has worked extensively on the position(s) of the verb in PIE. The model used is again the one presented in (5) and (6).

The unmarked order: verb-final

I start with unmarked verb-final in the AIELs and in PIE. This position of the verb can be found in many texts in AIELs at every turn, and it is unproblematic to reconstruct it for PIE. Already Delbrück (1967=1900: 83) wrote about PIE word order: “Das Verbum stand im unabhängigen Aussagesatz am Ende . . .” (In an independent declarative sentence the verb appeared at the end). One may refer to the following examples with verb-final taken from AIELs in this chapter: (4a–d), (13), (14), (15).

Now, in the texts of all AIELs one meets sentences where some linguistic material is put after the verb in the right periphery. Syntactic research by Gonda (1959) and McCone (1979) could clarify that much of this extraposed material falls into categories which are non-obligatory. Gonda (1959) calls sentences like that “amplified sentences” (cf. also Krisch 1997; 2001). One may discern the following types of extraposition (cf. Krisch 2001: 162–164): (a) apposition(s) to constituents overtly appearing in the VP before the verb; (b) extraposition of the subject (which is already expressed in the verbal form in AIELs and in PIE); (c) part of a Determiner phrase/Noun phrase, the other part being inside the VP (hyperbaton, see above); and (d) non-obligatory constituents. The following sentences are examples for these categories from Mycenaean Greek (25), Celtic (26), Vedic prose (27) and Hittite (28). The finite verb forms are in boldface.

## (25) Mycenaean Greek

a) apposition:

Ep 613,8 (Pylos):

[*sa-sa-wo*] *o-na-to*

*e-ke*

*ka-ma-e-u*

[Sasawo] beneficiary.right.ACC.SG. have.3SG.PRES. man.who.has.a.kama.NOM.SG.M.

‘Sasawo has a fief as Kamaeus.’

b) subject:

Le 641 (Knossos):

*o= a-po-te*

*de-ka-sa-to*

*a-re-i-jo*

so= from.afar receive.3SG.AOR.MID. Areios.NOM.SG.M.

‘So from outside Areios received (a delivery)’

c) part of an NP/DP:

Ep 704,3 (Pylos):

*e-ri-ta*

*i-je-re-ja*

*o-na-to*

*e-ke*

Eritha priest.NOM.SG.F. beneficiary.right.ACC.SG. have.3SG.PRES.

*ke-ke-me-na*

*ko-to-na*

*pa-ro*

*da-mo*

estate.GEN.SG. plot.of.land.GEN.SG. from community.DAT.SG.

‘Eritha, the priestess, has (got) a fief (consisting) of an estate, of a plot of land from the community.’

d) non-obligatory constituents:

Ma. 225,2 (Pylos):

*o-da-a<sub>2</sub>*

*ka-ke-we*

*o-u= di-do-si*

*za-we-te*

and.so bronze.smith.NOM.PL. not= give.3PL.PRES. this.year

‘and so the bronze-smiths do not deliver this year’

(26) Celtic

a) apposition:

Celtiberian: Botorrita 95–99 (Meid 1993)

*Tocoitei ios urantiom= ue auseti*

Tocoit.DAT.-LOC.SG. who.NOM.SG. pasture.ACC.SG. or make.use.of.3SG.AOR.SUBJ.

*aratim= ue, . . .*

arable.land.ACC.SG. or . . .

‘Who wants to use pastureland or arable land in the Togoit-(district) (Togoit is the name of a Celtiberian god, TK), . . .’

b) subject:

Gaulish (Genouilly, Stele) (after Lambert 1994: 94)

*ELVONTIV IEVRV. ANEVNO OCLICNO. LVGVRIX*

Elontios.DAT.SG. make.3SG.PRET. Aneunos.NOM.SG. Oculos.ADJ.NOM.SG. Lugurix.NOM.SG.

*ANEVNICNO*

Aneunos.ADJ.NOM.SG.

‘Aneunos, (son) of Oculos, (and) Lugurix (son) of Aneunos have made (3SG.PRET.!

TK) (me) for Eluontios (name of a god? TK)

c) part of a DP/NP

Celtiberian: Botorrita 23–25 (Meid 1993)

*uta oścues śtena uersoniti, śilapur*

and whoever.NOM.SG. this.ACC.PL.N. carry.through.3SG.PRES. silver.ACC.SG.

*śleitom conścilitom capiseti, cantom*

śleitom.ACC.SG. conścilitom.ACC.SG. take.3SG.AOR.SUBJ. 100

*śancilištara otanaum tocoitei eni. eni.*

śancilištara.ACC.PL.N. giving.to.ACC.SG. Togoit.DAT.-LOC.SG. into

‘And whoever carries this through, will take śleitom- and conścilitom-silver, (namely) 100 śancilištara (monetary unit, TK), for the giving-to into the (temple of) Togoit-’

d) non-obligatory constituents

Celtiberian: Botorrita 23–35 (Meid 1993)

*uta oścues śtena uersoniti, śilapur*

and whoever.NOM.SG. this.ACC.PL.N. carry.through.3SG.PRES. silver.ACC.SG.

*śleitom conścilitom capiseti, cantom*

śleitom.ACC.SG. conścilitom.ACC.SG. take.3SG.AOR.SUBJ. 100

*śancilištara otanaum tocoitei eni.*

śancilištara.ACC.PL. giving.to.ACC.SG. Togoit.DAT.-LOC.SG. into

‘And whoever carries this through, will take śleitom- and conścilitom-silver, (namely) 100 śancilištara (monetary unit, TK), for the giving-to into the (temple of) Togoit-’

(27) Vedic

a) apposition:

JB 1,28

*tam eva tābhir āhutibhiś śamayitvā ıṛjam*

this.ACC.SG.M. so this.INSTR.PL.F. rite.INSTR.PL.F. appease.ABS. power.ACC.SG.

*lokānām jayati yamam devam devānām*

world.GEN.PL. win.3.SG.PRES. Yama.ACC.SG. god.ACC.SG. god.GEN.PL.

‘having appeased him, the god of the gods Yama with these rites, he wins the power of the worlds.’

b) subject (with an emendment by Caland 1970=1919: 27, fn. 1)

JB 1,117,1:

*etasya ha vā idaṃ sāmnaḥ kṛte varṣati*  
 This.GEN.SG.N. PTCL. PTCL. present song.GEN.SG.N. for.the.sake.of rain.3SG.PRES.  
*ca parjanya uc ca grhṇāti.*  
 and Parjanya.NOM.SG. out and seize

‘for the sake of this present song (god) Parjanya lets it rain and stops (raining).’

c) part of a DP/NP

AB 2,14,6:

*amṛtam vā ājyam, amṛtam hiraṇyam*  
 Ambrosia.NOM.SG. PTCL. butter.NOM.SG. ambrosia.NOM.SG. gold.NOM.SG.  
*tatra sa kāmā upāpto ya ājve.*  
 there he longing.LOC.SG.M. arrive.PERF.PTCP. which.NOM.SG.M. butter.LOC.SG.  
*tatra sa kāmā upāpto yo hiraṇye.*  
 there he longing.LOC.SG.M. arrive.PERF.PTCP. which.NOM.SG.M. gold.LOC.SG.  
 ‘Butter is ambrosia, gold is ambrosia. There he has arrived in the longing which (is) in butter, there he has arrived in the longing which is in gold.’

d) non-obligatory constituents

AB 2,11,1:

*te devāḥ pratibudhyā ḡnimayīḥ puras*  
 this.NOM.PL. god.NOM.PL. perceiving.ABS. fiery.ACC.PL. citadel.ACC.PL.  
*tripuram paryāsyanta yajñasya ca*  
 triple.fortification.ACC.SG. place.around.3PL.IMP.F.MID. sacrifice.GEN.SG. and  
*āmanas ca gupṭyai*  
 self.GEN.SG. and preserving.DAT.SG.  
 ‘The gods, perceiving, placed around fiery citadels, a triple fortification *for the preservation of the sacrifice and of themselves.*’

(28) Hittite

a) apposition

Ges. 2, §86a, 27–29 (Friedrich 1971: 86)

(27) *takku DUMU-an an[nanumanzi]* (28) *kuiški pāi naššu*  
 If son.ACC.SG. be.educated.INF. anybody give.3SG.PRES. either  
<sup>LÚ</sup>NAGAR n[(ašma <sup>LÚ</sup>S[IMUGA <sup>LÚ</sup>UŠ.B]AR (29) *našma <sup>LÚ</sup>AŠGAB našma*  
 carpenter or smith weaver or leather.worker or  
<sup>LÚ</sup>TÚG . . .  
 fuller . . .

‘if anybody gives a son to be educated, either as a carpenter or as a smith, weaver or as a leather worker or as a fuller . . .’

b) subject

Ges. II § 76, 29 (Friedrich 1971: 82)

*tákku=ššan GEDIM-it tianzi LÚ-MA*  
 if= PTCL. ghost.of.a.dead.person.INSTR. step.forward.3PL.PRES. man=alas  
 MUNUS<sup>TUM</sup>

woman.NOM.SG.

‘if they step forward, man – alas- (and) woman.’

c) part of a DP/NP

StBoT 8, Vs I, 21’–25’ (cf. Otten-Souček 1969: 20)

(21) *apūš hantezumni tēḫḫi, úeš=a namma anda* (22)  
 this.ACC.PL. forecourt.DAT./LOC.SG. put.1SG.PRES. we=CoT again inside

[p]aiwani II <sup>D</sup>hantašepuš ḫarwani GI[(Š-aš)]  
 go.1PL.PRES. two Hantasepa-deities.ACC.PL. have.1PL.PRES. wood.GEN.SG.  
 ‘I put these (unclear reference in the text, TK) into the forecourt and we go inside again. We have two Hantasepa-deities (made) of wood.’

d) non-obligatory constituents

StBoT 8,VS I, 14’-15’ (cf. Otten-Souček 1969: 18)

DUMU.É.[(G)]AL šuppi wātar parā ʔpzi LUGAL-i  
 palace-official clean.ACC.SG.N. water.ACC.SG.N. out hold.3SG.PRES. king.DAT.SG.  
 MUNUS.LUGAL=ya  
 queen=and

‘The palace-official holds out clean water for the king and the queen’

(28c), part of which has already been cited as a hyperbaton construction (11), shows the amplified construction in the context of a number of regular Hittite OV sentences.

Marked order 1: verb-initial (V-1)

As we have seen, the PIE base-type was OV, a type which also encompasses the “amplified” structures in 2.1.3.1.

Under certain conditions, a movement of the verb to the top of the sentence was allowed. Whereas, e.g., in NHG this move is connected with certain sentence types (yes/no questions and imperative sentences show the finite verb at the beginning of the sentence), the AIELs and therefore PIE had a pragmatic condition for putting the verb at the top (V-1) (see below).

In the beginning of this paper we reconstructed word order variation (VO vs. OV) for PIE, leading to the results in (29a) and (29b):

(29) a) (=1)) (based on Vedic, Avestan and Greek)

PIE \*(é)- g<sup>wh</sup>en-t \*(H)e/og<sup>wh</sup>-im  
 (Augment)-kill-3SG.INJ.PRES./(IMPF.) serpent-ACC.SG.  
 →  
 ‘He kills /(killed) the serpent’

b) (=3)) (based on Vedic, Greek and Hittite)

PIE \*(H)e/og<sup>wh</sup>-im \*(é)- g<sup>wh</sup>en-t  
 serpent-ACC.SG. (Augment)-kill-3SG.INJ.PRES./(IMPF.)  
 ←  
 ‘He kills /(killed) the serpent’

Much has been written about the reasons for fronting the verb in AIEL and in PIE, e.g.

- Delbrück (1967=1900: 81) (my translation): “a word [moves] to the front or to the front edge of the sentence respectively, if it is of paramount importance for the statement . . . A verb can be important, if, in an isolated sentence, the action is more important than the other components . . . [I]n a series of sentences, it pushes the narrative or the discussion ahead.”
- Dressler (1969: 3) (my translation): “Verb-first is textually cataphoric or (much more frequent) anaphoric and (as a word order variant) characterizes textually bound sentences.”

- Dressler (1971: 18) (my translation): “Verb-first points to another sentence (either before or after) to which it is connected.”
- Klein in his article on Vedic verb-first states that the fronting of constituents seems universally to be associated with salience (Klein 1991: 135). He lists some environments where one encounters verb-first in Vedic texts: in imperative sentences, in iterative anaphoric structures, in chiasmic constructions, in cases of simultaneity between the event referred to in the hymn and actions occurring at the same time within the ritual, in the situation of fronting of verbs of speaking before a quotation, and in case of the Vedic verb *han-* ‘smite’ a tendency to front it.
- Luraghi (1995: 380): “On account of its lower frequency, the verb initial order was in a certain sense ‘abnormal’. By an iconic principle, it came to be used in cases where something abnormal was going on, either in the arrangement of the discourse, or in the course of the events reported.”
- Krisch (2004: 116): “Proto-Indo-European functions: . . . V-1: establishing a very close connection between sentences /sequences of action.”
- Bauer (2011: 47) (about Hittite; my translation): “Thus, one can note that it is the distribution of topic and comment in the sentence which is crucial for V1. The verb is the dummy-topic and the comment is the focus.”

Depending on the text, there is some truth in all of these statements. In my present view,

- verb-first introduces an important (often unexpected new) topic for the narrative with the consequence that further comments are given on it. In a coherent narrative, its general function is to accelerate narration, to drive forward the action.

Let us look again at (29a)=(1) in a fuller context, namely stanzas 1–3 of our Vedic example (2a):

(30) RV 1,32,1–3

1. <i>índrasya</i>	<i>nú</i>	<i>vīryāni</i>	<i>prá</i>	<i>vocam,</i>	<i>yāni</i>
Indra.GEN.	now	manly.deeds.ACC.PL.	PRV.	proclaim.1SG.INJ.AOR.	RP
<i>cakāra</i>	<i>prathamāni</i>	<i>vajrī</i>		<i>áhann</i>	<i>áhim,</i>
make.3SG.PERF.	first.ACC.PL.	thunderbolt-wielder.NOM.SG.		kill.3SG.IMPF.	serpent.ACC.SG.
<i>án<sub>u</sub></i>	<i>apás</i>	<i>tatar<sub>da</sub>,</i>	<i>prá</i>	<i>vakṣānā</i>	<i>abhinat</i>
PRV.	water.ACC.PL.	disclose.3SG.PERF.	PRV.	cave.ACC.PL.	cleave.3SG.IMPF.
<i>párvatānām</i>					
mountain.GEN.PL.					

‘I now proclaim the manly deeds of Indra, the first ones that he achieved, the thunderbolt-wielder. He killed the dragon, he disclosed the waters, he cleft the caves of the mountains.’

2. <i>áhann</i>	<i>áhim</i>	<i>párvate</i>	<i>śísriyānām.</i>	
kill.3SG.IMPF.	serpent.ACC.SG.	mountain.LOC.SG.	lean.PERF.PTCP.MID.ACC.SG.	
<i>tvāṣṭā</i>	<i>śmai</i>	<i>vájraṃ</i>	<i>svarīyaṃ</i>	<i>tataḥṣa</i>
Tvaṣṭar	he.DAT.	thunderbolt.ACC.SG.	shining.ACC.SG.	produce.3SG.PERF.
<i>vāśrā</i>	<i>iva</i>	<i>dhenávaḥ</i>	<i>syándamānā</i>	
mooring.NOM.PL.	like	cow.NOM.PL.	hurrying.up.PRES.PTCP.MID.NOM.PL.	
<i>áñjah</i>	<i>samudrām</i>	<i>áva</i>	<i>jagmur</i>	<i>ápaḥ</i>
straight	confluence.ACC.SG.	PRV.	go.3PL.PERF.	water.NOM.PL.

‘He killed the dragon who was lying on the mountain. God Tvaṣṭar has produced a shining thunderbolt for him. Like the mooing cows (go to their calves) hurrying up, the waters glided straight downwards to the confluence.’

3. <i>vṛṣāyámāṇo</i>		<i>ṛṇīta</i>		<i>sómaṃ</i>
Behave.like.a.bull.PTCP.PRES.MID.NOM.SG.		choose.3SG.IMP.F.		Soma.ACC.SG.
<i>trikadrúkeṣu</i>	<i>apibat</i>	<i>sutásya</i>		
trikadrúka-day.LOC.PL.	drink.3SG.IMP.F.	pressed.one.GEN.SG.		
<i>á sáyakam</i>	<i>maghávā</i>	<i>á datta</i>		<i>vájram.</i>
PRV. missile.ACC.SG.	munificent.one.NOM.SG.	take.3SG.IMP.F.MID.		thunderbolt.ACC.SG.
<b>áhann</b>	<b>enam</b>	<b>prathamajám</b>	<b>áhinām</b>	
kill.3SG.IMP.F.	he.ACC.	first-born.ACC.	serpent.GEN.PL.	

‘The one behaving like a bull (i.e. god Indra, TK) chose Soma, in the Trikadrúka-days he drank from the pressed one. The munificent one took the missile, the thunderbolt. He killed him, the first-born of the dragons.’

Note that the VO order with the lexical material of ‘kill’ and ‘snake’ reoccurs another two times in the three stanzas in (30). In all these cases the order is VO and not OV. Example (30) 1 starts with an OV sentence with the verb form *vocam* in absolute final position. Then follows the most important action of the god Indra, that he slew the dragon Vṛtra, a verb-initial sequence (starting with *áhann*). That Indra disclosed the waters is a consequence of his slaying of the dragon who locked them up. Also, the cleaving of the mountain serves to get the waters free. These actions, since they are not so important, are presented in a sentence which is absolutely verb-final (*tatarda*) and in an OV sentence with amplification (genitive *párvatānām* belonging to *vakṣānā*). In (30) 2, the important fact of killing Vṛtra is again emphasized by a verb-initial sentence (starting with *áhann*). In the following context, the scene is broadening again and describes things preceding the killing (that a thunderbolt was made for him) and, again, the situation afterwards (the waters glide downwards). (30) 3 goes back to the fighting scene and describes how the god Indra prepared himself to fight: He drinks soma and takes his weapon. Then the main action of Indra, the killing of the dragon, is repeated with the verb *áhann* in initial position.

There is no coherent narration; aspects of the topic are presented in little scenes jumping forward and backward. This function of comments, namely subdividing the scene opened by *áhann* ‘he kills’, is chosen in the whole hymn RV 1,32, where subdivisions of the act of killing are brought into little scenes with repetition. This hymn is like a musical piece consisting of a theme plus variations.

In contrast to this VO order, the OV order of the sentence dealing with Indra and the dragon, cited again in (31), does not indicate the TOPIC of what follows. It is part of new information given in the hymn, thus belonging to the informational focus of the text, the comment.

(31) = (4b) Vedic RV 2,15,1 (*áhim* < PIE \*(H)e/og<sup>wh</sup>-im; *jaghāna* < PIE \*g<sup>wh</sup>e-g<sup>wh</sup>on-e)

<i>trikadrúkeṣu</i>	<i>apibat</i>	<i>sutasya</i>	<i>asyá</i>	<i>máde</i>
Tríkadrúka-day.LOC.PL.	drink.3SG.IMP.F.	Soma.GEN.SG.M.	he.GEN.SG.	intoxication.LOC.SG.
<i>áhim</i>	<i>índro</i>	<i>jaghāna:</i>		
serpent.ACC.SG.	Indra.NOM.SG.	slay.3SG.PERF.		



‘At the Trikadrúka-days he drank from Soma; in the intoxication of / (caused by) this (one), Indra has slain the serpent’

The hymn RV 2,15 contains a short enumeration of all the noteworthy deeds of Indra without going into details. To illustrate this fact, (32) cites the first three stanzas of this hymn, which consists of ten stanzas:

(32) RV 2,15 (translation Griffith (1896); bold passage marked by TK)

1. Now, verily, will I declare the exploits, mighty and true, of him the True and Mighty.  
In the Trikadrukas he drank the Soma then in its rapture **Indra slew the Dragon**.
2. High heaven unsupported in space he established: he filled the two worlds and the air's mid-region.  
Earth he upheld, and gave it wide expansion. These things did Indra in the Soma's rapture.
3. From front, as 'twere a house, he ruled and measured; pierced with his bolt the fountains of the rivers, And made them flow at ease by paths far-reaching, These things did Indra in the Soma's rapture.

In this hymn the scene that Indra slew the dragon appears only once, namely in stanza 1, cf. the passage given in boldface. In the other verses various deeds of Indra are mentioned. Even stanza 3, which one could associate with the myth of the dragon, is formulated in a neutral way without any overt hints toward the dragon-story.

This enumeration of the deeds of Indra is underlined by a refrain starting with stanza 2 and going until stanza 9: "These things did Indra in the Soma's rapture."

Let us now turn to the Pindar examples (2c–d above) in Greek, which are interesting because they contain a native speaker's, namely Pindar's or, more correctly, the narrator's, comment on how to interpret the order with the verb in the front. The example for VO from the Pythian Ode Nr. 4, which has been mentioned earlier, is repeated as (33) (κτεῖνε ὄφιν):

(33) = (2d)

Pindar, Pyth. 4,249: [κτεῖνε < PIE \*tkenye°/kpenye° (injunctive present) or < PIE \*tken-s°/kpen-s° (injunctive aorist); ὄφιν < PIE \*(H)e/og<sup>wh</sup>-im]

κτεῖνε . . . ὄφιν  
kill.3SG.IMPF./AOR. serpent.ACC.SG.  
'he killed the serpent /dragon'

This text is about the Argonauts and the winning of the golden fleece. Jason the Argonaut has to get the golden fleece, which lies in a sacred grove in Kolchis. In (34), the context of (33) is offered in an English translation:

(34) Pindar, P. 4, 240–252 (translation, see: Odes, translated by Diane Arnson Svarlien [The Annenberg CPB Project, the Perseus Digital Library, accessed 9 September 2014])

But at once the marvellous son of Helios spoke of the shining fleece, telling where the sword of Phrixus had stretched it out. He expected that Jason would not be able to accomplish this further labour. For the fleece lay in a thicket, held in the ravening jaws of a serpent, [245] which in thickness and length surpassed a ship with fifty

oars, built by the blows of a hammer. It is too long a way for me to go by the beaten track; for time presses, and I know a shortcut. In poetic skill I am a guide to many others. Jason killed the grey-eyed serpent with its dappled back by cunning, [250] Arcesilas, and stole away Medea, with her own help, to be the death of Pelias. And they reached the expanses of Ocean, and the Red Sea, and the race of the Lemnian women, who killed their husbands.”

There is a remarkable breaking-off in the middle of this passage, where Pindar intervenes as a poet and reflects on his technique of presentation. These verses interrupt the calm flow of narration, and they are, on the one hand, a rather boastful demand the narrator makes of himself to speed up, and, on the other hand, they characterize the following lines as a short, condensed, so to speak “fast flowing,” version of the narrative. The original text of this passage is cited in (35).

(35) Pindar P 4, 247–248

μακρά μοι νεῖσθαι κατ’ ἀμαξίτον· ὥρα γάρ  
long.ADV I.DAT. go.INF.MID. along beaten.track.ACC. time.NOM. for  
συνάπτει καί τινα οἶμον ἴσαμι βραχύν·  
be.nigh.at.hand.3SG.PRES. and some way.ACC. know.1SG.PRES.ACT. short.ACC.  
πολλοῖσι δ’ ἄγημαι σοφίας ἐτέροις.  
many.DAT.PL. but lead.1SG.MID.PERF. skill.ACC.PL. other.DAT.PL.

‘It is too long a way for me to go by the beaten track; for time presses, and I know a shortcut. In poetic skill I am a guide to many others.’

Pindar’s remarks in (35) offer a rather unique native speaker’s intuition about the functions of VO in quite a clear way. The next sentence, our sentence (33)=(2d), starts with marked verb-first, showing the poetic skills of Pindar, which he himself just mentioned. The VO sentences that follow (35) are cited in (36) in full. The two verbs in sentence-initial position are put in bold:

(36) (cf. (33)) Pindar P 4, 249–250:

**κτεῖνε** μὲν γλαυκῶπα τέχνας ποικιλόνωτον ὄφιν,  
Kill.3SG.IMPF. PTCL. grey-eyed.ACC. cunning.DAT.PL. dappled.back.ACC. serpent.ACC.  
ὃ Ἀρκεσίλα, **κλέψεν** τε Μήδειαν σὺν αὐτᾷ τὴν  
o Arkesilaos.VOC. steal.AOR.3SG. and Medea.ACC.SG.F. with she.DAT. she.ACC.SG.F.  
Πελίαο φονόν· ἔν τ’ Ὀκεανοῦ πελάγεσσι  
Pelias.GEN. death.ACC. PRV. and Ocean.GEN. flood.DAT.PL.  
μίγεν  
make.acquaintance.with.3PL.AORII

‘He (i.e. Jason, TK) **killed** the gray-eyed serpent with its dappled back by cunning, [250] Arcesilas, and **stole away** Medea, with her own help, to be the death of Pelias. And they reached the expanses of Ocean (the name of a god and a river which is the origin of all oceans, TK) . . .’

Basically, the function of verb-first in (36) is similar to the one which we have seen in Vedic: The scene is opened up by the verb, which states an important topic pushed forward in a lively way and leads to further comments: the action is pushed ahead by a second verb form in initial position (verb κλέψεν). With the last verb-final sentence (verb μίγεν), the scene comes to a rest.



On the other hand, the Pindar citation (37)=(4d), the evidence for verb-final position, has a context similar to the comparable Vedic sentences with the verb in final position ((31) and (32)): The deeds of the hero Bellerophontes are simply enumerated and not expanded; they do not form the basis for a narration. The killing of the dragons is part of this enumeration. The context of (37) is cited in (38).

(37)=(4d) Greek: Pindar, O. 13,88 ff. (ἔπεφνε(v) < PIE \*e-g<sup>wh</sup>e-g<sup>wh</sup>n-e reduplicated aorist)

καὶ Χίμαιραν πῦρ πνέοισαν καὶ Σολύμους ἔπεφνε

καὶ Χίμαιραν πῦρ πνέοισαν καὶ  
and Chimaira.ACC.SG.F. fire.ACC.SG.F. breathe.PTCP.AOR.ACC.SG.F. and  
Σολύμους ἔπεφνε  
Solymoi.ACC.PL. kill.3SG.AOR.

‘And he (i.e. Bellerophontes, TK) killed the fire-breathing Chimaira and the Solymoi’

(38) Pindar, O. 13,88 ff. (translation based on Odes, translated by Diane Arnson Svarlien [The Annenberg CPB Project, the Perseus Digital Library, accessed 9 September 2014])

σὺν δὲ κείνῳ καὶ ποτ’ Ἀμαζονίδων αἰθέρος  
with but that.one.DAT.SG. and once Amazones.GEN.PL. heaven.GEN.SG.  
ψυχρῶν ἀπὸ κόλπων ἐρήμου τοξόταν  
cold.GEN.PL. from bosom.GEN.PL. lonely.GEN.SG. archer.army.ACC.SG.  
βάλλων γυναικεῖον στρατόν καὶ  
attack.PTCP.PRES.NOM.SG.M. female.ACC.SG. army.ACC.SG. and  
**Χίμαιραν πῦρ**  
Chimaira.ACC.SG.F. fire.ACC.SG.F.  
πνέοισαν καὶ **Σολύμους ἔπεφνε.**  
breathe.PTCP.AOR.ACC.SG.F. and Solymoi.ACC.PL. kill.3SG.AOR.

‘And with that one (i.e. Pegasus, TK), from the chilly bosom of the lonely air, he (i.e. Bellerophontes, TK) once, – attacking the Amazons, the female army of archers, – killed the fire-breathing Chimaera, and the Solymi.’

From these and similar examples one may hypothesize that Greek and Vedic, and presumably also IE, had both possible verb orders, verb-initial and verb-final, and that the difference between them is that the final position of the verb shows a regular sequence with the topic in front of it and the comment in the predicate part, whereas the verb-initial construction serves stylistic purposes indicating a (new) topic, represented by the action expressed by the verb and its object, which leads to further comments directly derived by the verbal action.

In my model (see (5) and (6)), sentences with an initial finite verb can be represented by movement of the finite verb into the C position. This is a sort of movement which is well established in generative grammar (cf., e.g., Haegeman 1994: 302). Since Vedic (like PIE) is a fully inflecting language, the empty subject *pro* is licensed by the agreement morphology present in *āhann* (cf. also Axel 2007: 314 and Sternefeld 2006: 613, referring to Barbosa 1995). A syntactic tree of part of the Vedic sentence in (30) 2 is (39):

(39)

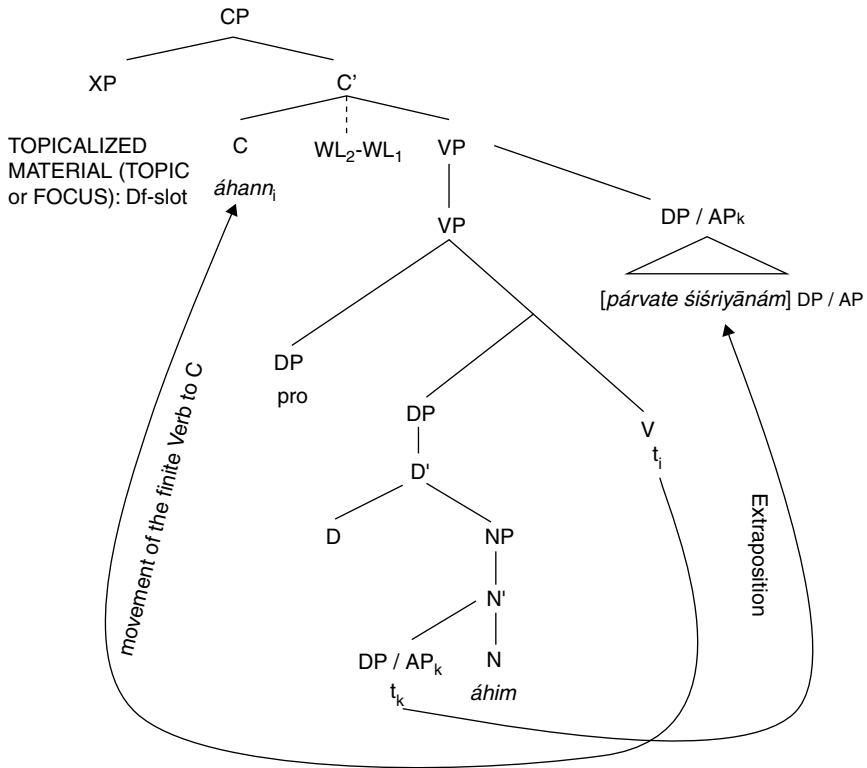


FIGURE 1.7 STRUCTURE OF RIGVEDA 1,32,2

For a Greek instance of V-1 cf. example (41).

Marked order 2: verb-second (V-2) / verb-third (V-3)

Whereas the existence of verb-first in PIE is uncontroversial, the reconstruction of verb-second (as in modern German) can be found mainly in the work of the author of this article. But also Keydana (2011: 109–110) assumes verb-second at least for Vedic. The ideas presented here are work in progress.

In fact, verb-second is structurally identical to verb-initial with **scheme 1** in the approach used here (cf. (5)): The verb is moved into C. The only difference is that (in the case of verb-initial, cf. (36)) the Df-slot is empty, whereas in verb-second the Df-position is filled. If one takes **scheme 2**, V-2 is possible if the Df-slot is filled, the verb is moved to C, and the second XP in front of the C position is empty. With scheme 2, verb-third (V-3) is a possible option when the verb is moved into C and the two XP positions in front of it are also filled. According to Krisch 2004 the function of verb-second in PIE is similar to

the function of verb-initial. Its function is “establishing a connection between sentences/sequences of action, but not as close as verb initial” (Krisch 2004: 116). It is viewed as a probably late IE feature, since this phenomenon can be observed very frequently in Ancient Greek and in the Germanic languages, whereas it is rare in Vedic and seems to be absent in Hittite (cf. Krisch 2004: 119). Whether verb-third has a special function is still unclear but it probably does not differ much from unmarked V-end.

A clear instance of V-3 is Il. 18, 468 (*μὲν* is interpreted as a WL3 particle (see section 2.1.2) which is attached directly to its host, *τὴν*). Since the local specification *αὐτοῦ* is demanded by the verb in my model, this cannot be an extraposed constituent.

(40) Il. 18,468

Ὦς εἰπὼν τὴν μὲν λῖπεν αὐτοῦ  
 thus speak.PTCP.AOR.NOM.SG.M. she.ACC.SG. verily leave.3SG.AOR. there  
 ‘After he had spoken thus, he (i.e. the god Hephasitos, TK) left her (i.e. Thetis, Achill’s mother, TK) there.’

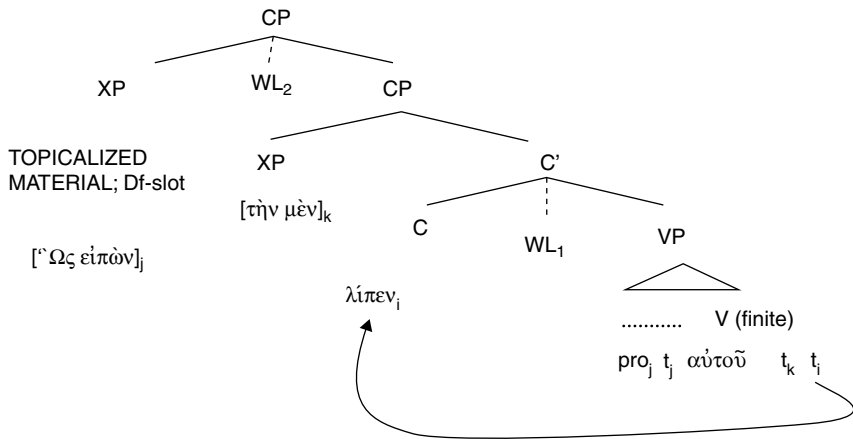


FIGURE 1.8 STRUCTURE OF HOMER, ILIAD 18,468 / 1

This sentence is part of the shield description of Achilles in the Iliad. As shown in Krisch 2001: 169–170, this sentence starts a sequence of sentences describing the lively scene where Hephaistos does his blacksmith’s job and one action follows the other like the V-1 sentence immediately following 40 in the text, cf. (41):

(41) Il 18, 468

βῆ δ’ ἐπὶ φύσας  
 go.3SG.AOR.INJ. but to bellow.ACC.PL.  
 ‘And he (i.e. the god Hephaistos, TK) went to the (pair of) bellows’

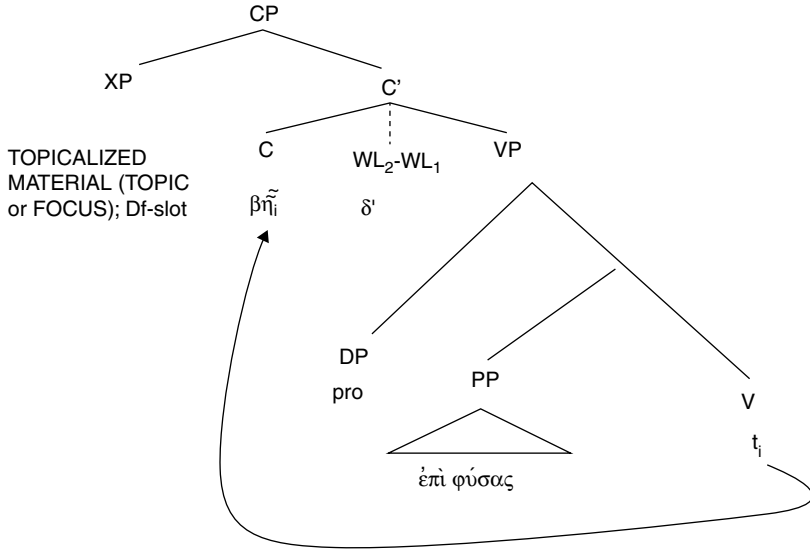


FIGURE 1.9 STRUCTURE OF HOMER, ILIAD 18,468 / 2

In the following illustration from the fifth book of Homer I want to show the phenomenon: that V-1 and V-2 appear in the same function (cf. also Krisch 1997: 299–300). But I also want to draw the reader’s attention to the fact that it is quite often not possible to decide formally whether a sentence is V-2 or verb-final, if the sentence is too short for a decision:

(42) Il. 5, 97–101

αἶψ'	ἐπὶ	Τυδεΐδῃ	<b>ἐπιταίνετο</b>	καμπύλα
quickly	against	Tydeus-son.DAT.	bend.3SG.IMP.F.MID.	curved.ACC.PL.N.
τόξα		καὶ βάλ'	ἐπαΐσσοντα	
bow.and.arrow.ACC.PL.N.		and hit.3SG.AOR.	assault.PTCP.PRES.ACC.SG.	
τυχῶν		κατὰ δεξιὸν	ὄμων	
succeed.PTCP.AOR.NOM.SG.M.		on right.ACC.SG.M.	shoulder.ACC.SG.M.	
θώρηκος	γύαλον	διὰ δ'	<b>ἔπτατο</b>	
corselet.GEN.SG.	hollow.ACC.SG.N.	through but	fly.3SG.AOR.MID.	
πικρὸς	οἴστός,	ἀντικρὺ	δὲ διέσχε,	
sharp.NOM.SG.M.	arrow.NOM.SG.M.	straight.through	but go.through.3SG.AOR.	
<b>παλάσσετο</b>	δ'	αἵματι	θώρηξ.	τῷ δ'
besprinkle.3SG.PASS.IMP.F.	but	blood.DAT.	corselet.NOM.SG.	then but

ἐπὶ μακρὸν ᾄσε Λυκάονος ἀγλαὸς υἱός.  
 thereupon loud.ADV. cry.out.3SG.AOR. Lykaon.GEN. splendid.NOM.SG. son.NOM.SG.  
 ‘Quickly he (i.e. Pandaros the son of Lykaon, TK) bent the curved bow and arrow against the son of Tydeus (i.e. Diomedes, TK) and hit the assaulting one successfully on the right shoulder, (on) the curvature of the corselet. And the arrow went straight through and the corselet was sprinkled with blood. And then, thereupon Lykaon the splendid son of Lykaon cried out.’

The first sentence in (42) is a V-3 sentence according to our model. The accusative *καμπύλα τόξα* cannot be an amplification, since it is an obligatory constituent. One can describe it with scheme 2 (cf. (6)):

(43)

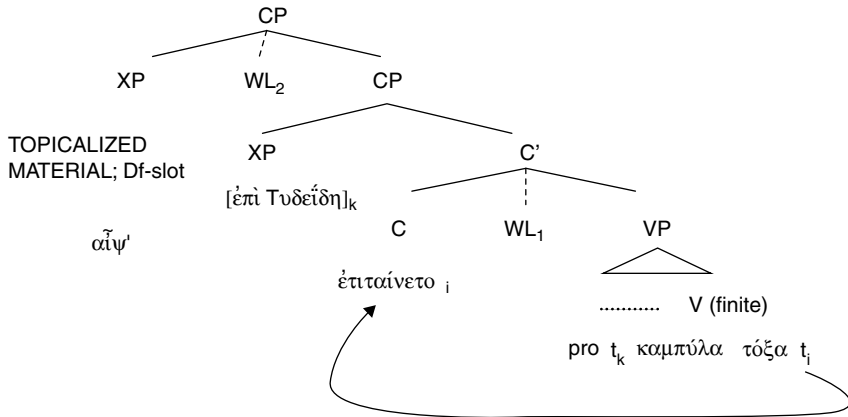


FIGURE 1.10 STRUCTURE OF HOMER, ILIAD 5,97

The next sentence (starting with *καὶ βάλ'*) is a V-1 sentence (like (39) and (41)), since the sentence connector *καὶ* does not “count”. The sentence starting with *διὰ δ' ἔπτατο* could be either a verb-final sentence or a V-2 sentence with *διὰ* in the Df-slot and *δ'* in WL<sub>2</sub>-position, again using scheme 2. (44) offers a tree for the interpretation as V-2:

(44)

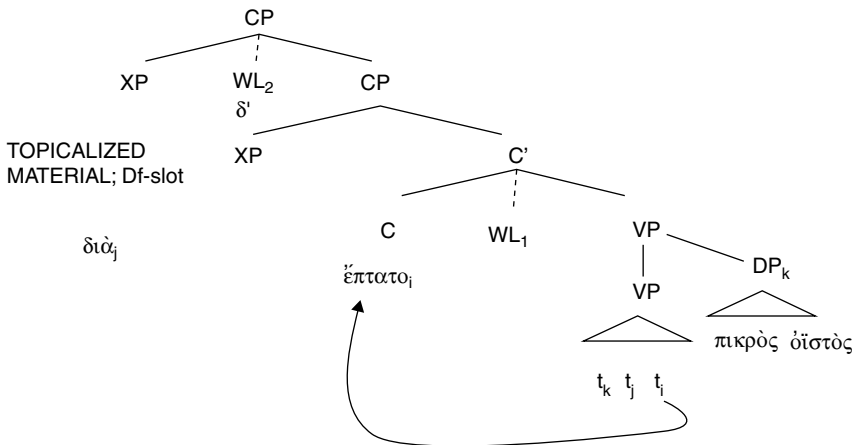


FIGURE 1.11 STRUCTURE OF HOMER, ILIAD 5,99

Also, the following short sentence (ἀντικρὺ δὲ διέσχε,) could be either an example for verb-final or a case of V-2. The tree for V-2 would look like (45):

(45)

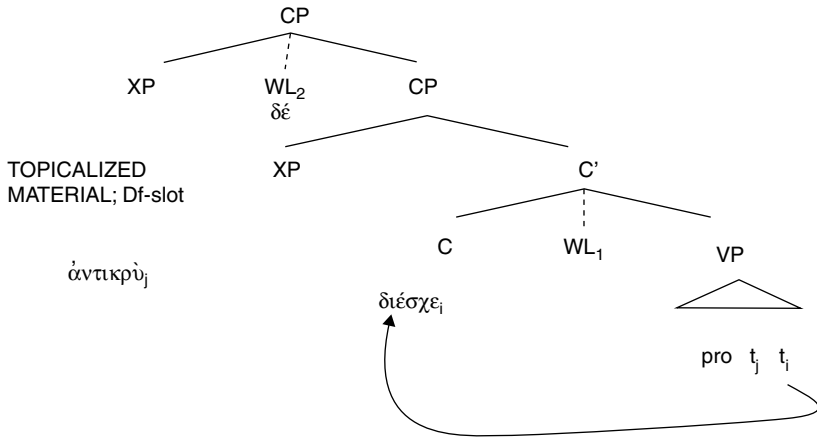


FIGURE 1.12 STRUCTURE OF HOMER ILIAD 5,100

The sentence starting with *παλάσσετο* is unambiguously V-1. The last sentence is unambiguously verb-final (with amplification); see (46):

(46)

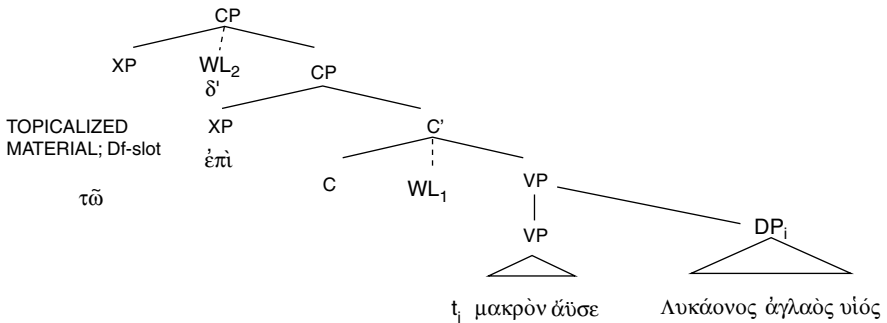


FIGURE 1.13 STRUCTURE OF HOMER, ILIAD 5,101

If our interpretations are right, the scene narrated starts with a V-3 sentence when Pandaros stretches his bow against Diomedes. The scene where the arrow hits Diomedes’ shoulder, penetrates the corselet and causes the blood to sprinkle is described by sentences which can be analysed as V-1 and V-2.

Conclusion:

- V-1 and V-2 have the function of driving forth the action in a narrative text, of making it lively. V-1 presumably is the “livelier” of the two.

*Object ellipsis*

A remarkable feature of PIE word order is object ellipsis. It is ubiquitous in Greek (cf. (47)), but also Latin, Vedic and Hittite show plenty of examples.

- (47) Ancient Greek Hom. Il 5, 22–24, cf. Krisch (2009: 207) (Dares, a priest of the god Hephaistos, has two sons, Phegeus and Idaïos. In the Trojan war, they fight against Diomedes. Diomedes kills Phegeus and would have killed Idaïos as well, but Hephaistos rescues Idaïos).

οὐδὲ γὰρ οὐδέ κεν αὐτὸς ὑπέκφυγε κῆρα μέλαιναν, / ἄλλ’ Ἥφαιστος ἔρυτο, σώωσε  
 δὲ νυκτὶ καλύψας, / ὥς δὴ οἱ μὴ πάγχυ γέρων ἀκαχήμενος εἶη. ‘Nay, he [Idaios, TK]  
 would himself not have escaped the black goddess of death, but Hephaistos guarded  
 HMM, rescued HMM, enfolding HMM in darkness so that the aged one [= Dares, TK]  
 might not be utterly fordone with grief’

ἀλλ’	ΜΗΝ	Ἥφαιστος	ἔρυτο,	σώωσε
but	HMM	Hephaistos	guard.3SG.AOR.MID.	rescue.3SG.AOR.
δὲ	ΜΗΝ	νυκτὶ	ΜΗΝ καλύψας	
and	HMM	night.LOC.SG.	HMM	hide.PTCP.AOR.NOM.SG.M.

After discussing examples from Latin, Vedic and Hittite and recent scholarly literature on the topic, Krisch (2009: 211) concludes that “PIE had a rule that allowed deleting objects which refer back to already mentioned ‘thematic’ material. This ‘null object’ type of ellipsis thus is operating in a forward direction.” Their frequency in coordinate conjunction is underlined by Keydana and Luraghi (2012).

**Absolute constructions**

Absolute constructions (henceforth ACs) of the type Latin *me absente* (I.ABL.SG. absent.PTCP. PRES.ABL.SG.) ‘in my absence’ or *piro florente* (pear.tree.ABL.SG. blossom.PTCP.PRES.ABL.SG.) ‘when the pear tree is in blossom’ may be found in a number of AIELs and are a candidate for reconstruction. Ruppel 2013 is the latest reference work on this topic with an extensive bibliography. According to her, ACs denote spaces or points in time with an obligatory qualifier (normally a participle/verbal adjective) when the respective head noun does not refer to time, but is “put into time” (Ruppel 2013: 31) by the predicatively used participle. She sees the IE origins in regular locative expressions of nouns denoting time, like “at dawn” and “in spring”, which, when applied to nouns not referring to time, are “put into time” by the qualifier. Thus, compare regular Vedic *uśási* ‘at dawn’ (dawn.LOC.SG., e.g. RV 4,2,8; expressing a definite point in time, conveyed by a locative) with *sūrye udyatí* (sun.LOC.SG. up-go.PTCP. PRES.LOC.SG.) ‘at sunrise’, where ‘sun’ is not per se expressing a point in time.

If one takes (as Ruppel 2013 does) the Vedic situation with the locativus absolutus as reflecting the PIE state, one may easily explain the Latin situation (with the ablative as the case form in the AC, which also encompasses the PIE locative via syncretism) or the Germanic situation (with a dative as the case form in the AC, which can also continue a PIE locative via syncretism). Greek, however, is problematic because the genitive which appears in the Greek AC (a genitivus absolutus construction) does not go back to a PIE locative, but can represent either a PIE ablative or genitive. One may argue here that genitives can be used as case forms for locative meaning in Greek also outside of ACs, e.g. *νυκτός* (night.GEN.SG.) ‘at night’, *πολέμω καὶ εἰρήνῳ* (war.GEN.SG. and peace.GEN.SG.) ‘in war and peace’ (Elis, 200 BC) (cf. Ruppel 2013: 220). Such genitives are traditionally classified as “partitive” in the handbooks (e.g. *νυκτός* ‘(time portion of) night’, *πολέμῳ καὶ εἰρήνῳ*

‘(time portion of) war and peace’). These constructions compete with constructions showing (preposition +) dative, cf. ἐν εἰρήνῃ καὶ ἐν πολέμῳ (in war.DAT.SG. and in peace.DAT.SG.; e.g. X.Lac. 11.1.2). The only attestation of this phrase in the dative which I could find without a preposition is in a Scholia in Iliadem 18, 490. The Greek dative can continue the PIE dative, instrumental and locative. Starting from such competing constructions, there could have been a transfer to the genitive. Unlike others (including Krisch 1988, Keydana 1997: 9), Ruppel is not interested in the semantic equivalence of ACs to finite subordinate clauses.

Krisch (1988: 8–9, following Holland 1986) reconstructs a *nominativus absolutus* (nominative + participle) for PIE. This construction is very rare but appears in a greater number of AIELs than the ACs with oblique cases. One can imagine these nominatives, loosely connected with the neighbouring sentence, as forming a frame for the adjacent sentence and can interpret (with Krisch 1988) the (presumably later, but still PIE) development of ACs with oblique case forms as a stronger integration of the temporal frames into the sentence. Since the locative with its temporal function is the ideal candidate for this, the rest of the story could roughly go as sketched above when dealing with Ruppel 2013. An example for a *nominativus absolutus* is (48):

(48) Greek, Hom. Il. 3,211:

ἄμφω	δ’	ἐζομένω,	γεραρότερος	ἦεν	Ὀδυσσεύς
both.NOM.DU.	but	sit.PTCP.MID.NOM.DU.	well-built.COMP.NOM.DU.	be.IMP.F.3SG.	Odysseus

‘when both (i.e. Menelaos and Odysseus, TK) were seated, Odysseus was the more impressive’

Other languages where this construction appears are Vedic, Latin, Germanic, Hittite and Lithuanian. The Hittite examples are doubtful, though (cf. Keydana 1997: 288).

An attempt to draw the semantic paths from an original temporal meaning of ACs to other meanings of ACs (conditional, causal, concessive, modal) is made by Krisch (1988: 9). Keydana (1997) offers a very extensive selection of attested ACs in AIELs arranged according to a number of important syntactic categories. Like the author of this chapter, he considers a late PIE origin for ACs (Keydana 1997: 34), but he does not take the temporal value of ACs as basic and wants to explain the relationship between the ACs and the rest of the sentence by relational semantics/pragmatics (Keydana 1997: 45).

## Case functions in the AIELs and in PIE

### *Introduction*

In IE studies, the semasiological approach has a long tradition. One takes the case form of a particular language and describes the functions which this case form can be associated with. This type of research is still ongoing and now centres on the question of determining the “core” meaning of the cases and their relation to more peripheral uses and on the question of which uses of the case in question can be taken over by other case forms. This type of research is surely still a rewarding task if it is carried out meticulously.

### *The instrumental*

The Vedic instrumental case

As an elaborate example for such an investigation, I mention Hettrich 2002, a specimen from a work in progress on R̥gvedic uses of case forms.<sup>4</sup> The article deals with the instrumental.



Hettrich (2002: 46) defines the core/prototypical meaning of the instrumental in the R̥gveda (RV) as the ‘instrumental of means’. According to him, an NP carrying the prototypical instrumental case exhibits the following features:

(49)

- a) It is -animate, concrete and easy to handle.
- b) It is a (physical) object the existence of which is independent of the actual situation and of the participants in this situation.
- c) This (physical) object is under the control of the one controlling the action and stays there during the action.
- d) This (physical) object enables the controller to carry out an action which is controlled by him or her, which is dynamic and oriented towards a patient, or facilitates such an action.

This prototypical usage is presented in (50):

(50) RV 2,15,6b:

*vájren<sub>a</sub> ána uśásaḥ sám pipeṣa*  
 Vajra.INSTR.SG. travelling.cart.ACC.SG. Uśas.GEN.SG. together demolish.3SG.PERF.  
 ‘With the Vajra (i.e. the thunderbolt, TK) he (i.e. the god Indra, TK) demolished the cart of Uśas (i.e. the goddess of dawn, TK)’

The other shades of meaning of the instrumental case are explained linguistically by the absence of one or more of the features of the prototypical usage enumerated above. If one takes away feature c) one gets sentences like (51):

(51) RV 2,10,4a:

*jígharmy agnīm haviṣā ghṛténa*  
 sprinkle.1SG.PRES. Agni.ACC.SG. oblation.INSTR.SG. butter.INSTR.SG.  
 ‘I sprinkle Agni with an oblation, with butter’

A further example is the causal meaning of an NP using an instrumental. This meaning is achieved if the NP slot is headed by a noun which is -concrete (i.e. +abstract), and it is sometimes doubtful whether there exists a full and sole control over it by the subject controlling the action. Consider example (52), where the subject is Indra, the mighty god who sets free the waters and makes springtime possible:

(52) RV 2,13,7 (Hettrich 2002: 49)

*yáḥ puṣpínīs ca pras<sub>a</sub>vaś ca dhármaṇā*  
 who.NOM.SG. blossoming.ACC.PL.F. and fruitful.ACC.PL.F. and divine.law.INSTR.SG.  
*ádhi dāne víy avánīr ádhārayaḥ*  
 onto.ADV. field.LOC.SG. asunder river.ACC.PL.F. bear.forth.2SG.IMPF.  
 ‘you who distributed the blossoming and fruitful (flowers) (and) the rivers onto the field according to/because of divine law’.

The causal usage of the instrumental in (52) corresponds to non-prototypical uses of the locative (53) and the accusative (54) (with the adverb *ánu*), and with a reading of the ablative (abstract source) which is fairly near the ablative prototype (55).

(53) RV 9,7,1:

*ásṛgram indavaḥ pathā dhármann*  
 release.3PL.AOR.PASS. sap.NOM.PL.M. path.INSTR.SG. divine.law.LOC.SG.

*ṛtāsya*      *suśrīyaḥ*  
 truth.GEN.SG.    very.beautiful.NOM.PL.M.  
 ‘The saps poured out along the path because of/according to divine law of truth, the  
 uttermost excellent (ones).’

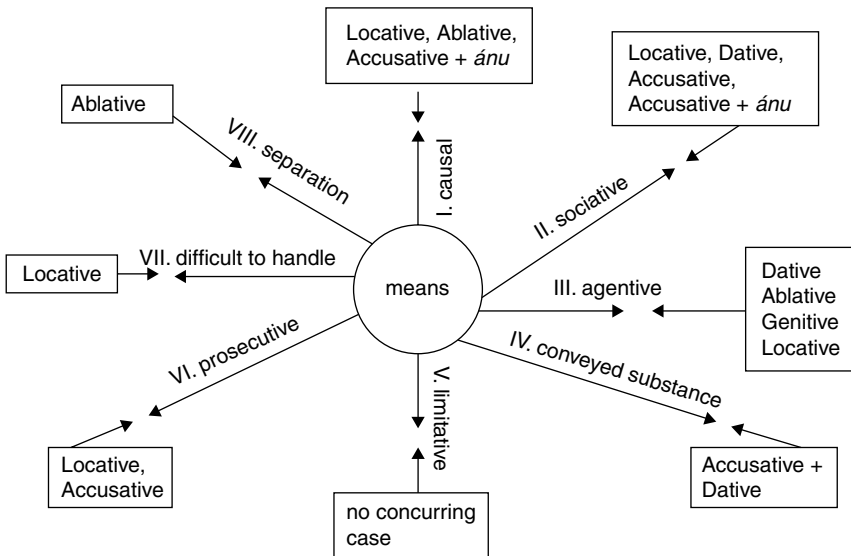
(54) RV 3,17,1:

*samidhyāmānaḥ*                      *prathamā*      *ānu*      *dhārmā*  
 kindled.PRES.PASS.PTCP.NOM.SG.    first.ACC.PL.N.    according.to    divine.law.ACC.PL.N.  
 ‘kindled according to the first divine laws (referring to Agni, TK)’

(55) RV 6,70,3:

*prá*      *prajābhir*      *jāyate*      *dhármaṇas*      *pári*  
 forth.ADV./PRV.    descendant.INSTR.PL.    create.3SG.PASS.    divine.law.ABL.SG.    around.ADV./PRV.  
 ‘he is created further through descendants because of/according to divine law  
 (being) around’

Hettrich uses the (slightly adapted and translated) diagram in Figure 1.14 to show the core and the peripheries of the usages of the instrumental in Vedic. This method of describing the function of cases in AIELs in my opinion is true progress compared to the traditional descriptions of the function of cases: It makes use of ideas from general linguistics (the use of features and the application of prototype theory), and it illustrates the results in a geometric symbolization (see Figure 1.14). Hettrich’s method is extendable to other languages and may lead to more detailed functional reconstructions. Any reconstruction is of course much less secure than the philological reasoning in a language with a clearly marked morphological instrumental.



**FIGURE 1.14** HETTRICH 2002: 47 (slightly adapted): THE SPECTRUM OF MEANING OF THE R̥GVEDIC INSTRUMENTAL (prototypical meaning in the circle, other meanings radiating with arrows) WITH CONCURRING CASES (inside the rectangles)

## The Hittite instrumental case and questions of reconstruction

Let us have a look at another AIEL, Hittite. We here make use of the very detailed and philologically excellent dissertation by Melchert 1977, dealing with the functions of the ablative and instrumental in all historical strata of Hittite.

Like in Vedic, the most common uses of the instrumental in Old Hittite (arguing functionally, not formally) are near its prototypical use, the instrumental of means. In (56) the prototype is evident:

- (56) (cf. Neu 1980, 105) Ritual text Nr. 43, Vs. I 12' (= KBo XVII 43 12'):  
*appa=ma=šta nēa* <sup>LÜ</sup>*mēnean* <sup>KUŠ</sup>*šarazzit* *walaḫzi*  
 back-but-PTCL. turn.3SG.MID. ritual.official.ACC.SG. waterhose.INSTR. beat.3SG.PRES.  
 ‘but he turns then, beats the ritual official with a waterhose.’

The instrumental <sup>KUŠ</sup>*šarazzit* ‘with a waterhose’ (the Hittite instrumental ending *-(i)t* is not a morphological cognate of any Vedic instrumental form, and its formal origin remains uncertain) fulfills all the criteria mentioned above for a prototypical instrument. As in Vedic, there are other uses, e.g. one where the feature c) (cf. (49)) is taken away:

- (57) Old Hittite, Zalpa-Text (cf. Otten 1973: 6) A Vs. 2:  
*tuppuš* *šakānda* *šunnaš*  
 container.ACC.PL. excrement.INSTR. fill.3SG.PAST  
 ‘she filled the containers with excrement’

Taking into account the Vedic and Old Hittite examples just mentioned, one may reconstruct the functions “prototypical instrumental” (cf. the examples (50) and (56)) and “prototypical instrumental use minus feature c)” (cf. the examples (51) and (57)) for PIE.

If one seeks for parallels to Vedic examples with a causal nuance like the one under (52) one does not find comparable uses in Hittite for the instrumental case, neither in Old Hittite nor in the subsequent stages of the language. This is a caveat for reconstructing this use of the instrumental case for PIE.

On the other hand, the ablative in Hittite (already in Old Hittite) can be found in causal contexts, just like the Vedic ablative (cf. (55) above):

- (58) IBoT I 36 I 58–59 (cited with translation after Melchert 1977: 289 f):  
*nu* <sup>LÜ</sup>NI-DU<sub>8</sub> *punuššanzi* *nu* <sup>GIŠ</sup>SUKUR.ḪI.A *apāš*  
 and doorkeeper interrogate.3PL.PRES. and spear.ACC.PL. he.NOM.SG.  
*naḫšaraz* (!) (i.e. *naḫšarattaz*) *uškizzi*  
 fear.ABL. see.3SG.PRES.  
 ‘They interrogate the doorkeeper and he will (always) see the spears out of fear.’

For a PIE reconstruction this means that the causal reading of the ablative is a candidate for reconstruction, whereas the causal reading of the instrumental seems to be an Indo-Aryan innovation.

Thus, an example like the Latin text from Plautus in (59) with a causal ablative will probably continue a PIE ablative and not a PIE instrumental:

(59) Plaut. Amph 1118

*nam mihi horror membra misero*  
 since 1SG.PRON.DAT. horror.NOM.SG. limb.ACC.PL. poor.DAT.SG.M.  
*percipit dictis tuis*  
 seize.3SG.PRES. word.ABL.PL. your.ABL.PL.

‘since because of your words horror seizes the limbs of miserable me’

### Case syncretism

“Case syncretism” is the linguistic term for a formal collapse of originally separate functions of cases, so that one case form covers different functions. In more recent times there have been attempts to make use of models to get insights into this phenomenon.

In his model of a causal order of participants, Croft (1991: 186) states that the (universal) causal chain of constituents in a sentence is influenced by a cause-result scheme, “cause” starting the causal organization of sentence and “result” ending it. Consequently, he distinguishes semantic roles which belong to the “antecedent” type (such as AGENT, COMITATIVE, MEANS, MANNER, INSTRUMENT, SOURCE and CAUSE) from semantic roles belonging to the “subsequent” type (such as DIRECTION, GOAL, RESULT, BENEFACTIVE, RECIPIENT and LOCATIVE). Croft claims that case syncretism is preferably possible inside these types, viz. that case syncretism is preferred among antecedent roles or among subsequent roles, respectively. Croft offers a table of 35 languages surveyed (Croft 1991: 237–239). In principle, he takes into account synthetic case forms and adpositional phrases, but he does not discuss data from AIELs in any detail.

If we interpret the syncretism of the PIE ablative and genitive into the Greek genitive (cf. also Luraghi 1987: 362–363) along these lines, one could argue that the main function of the genitive case (also in PIE) is the marking of a (nominal) dependency. In the case of the frequent genitivus subjectivus (e.g. Gr. φόβος τῶν πολεμίων ‘fear of the enemies’ (meaning: ‘the enemies fear s.b.’) the genitive denotes an Agent and belongs to the “antecedent” group, just like the ablative with its core semantics (Source).

It appears, though, that this model is too simplistic for the description of the syncretisms observed in AIELs. Luraghi (2001) argues that “illicit” syncretisms across Croft’s two groups do occur. Her main point is that not only semantics has to be considered when dealing with syncretism but also syntax and the lexicon. This point was already made by Serbat (1992: 281), who deals with the syncretism of the PIE ablative, instrumental and locative (note: for Croft, the locative belongs to a different group than the ablative and instrumental) into a single case called “ablative” in Latin. He emphasizes that this syncretism has to be interpreted as mainly syntactic, since these cases form the central domain of the circumstantial cases (“adjuncts”), whereas the rest of the cases represent the core syntactic relations (“arguments” in the narrow sense). The Latin syncretism thus underlines a syntactic opposition formally. But one has to be careful about making too general statements: The ablative in Latin is also a formal mixture: The consonant stems still preserve the locative ending as the “ablative” in the singular, and the pure “ablative” in Latin was used mainly to express instrumental meanings, while the other functions were expressed either by certain prepositions + ablative or by lexical restrictions on certain verbs (cf., e.g., the pure “ablative” designating a location with *incolit* ‘lives (in)’ (Pl. Rud. 907)). In her analysis of the syncretism of the PIE dative,

instrumental and locative into the Greek dative, Luraghi (2001: 42) emphasizes that lexical features like +/–animate are important: If the Greek dative reflects a PIE dative, it has to have the feature [+animate]; if it reflects the PIE instrumental, the relevant feature is [–animate]; and in the case of an original PIE locative, the following lexical features have to co-occur: [–animate], [+place] / [+time]. In the exceptional case where someone who is [+animate] is used as an instrument, a prepositional phrase *διὰ* ‘through’ + genitive is used (Luraghi 2001: 42).

### Argument structure

This section, though it deals with one of the most discussed issues in PIE syntax, will be kept very short here, due to lack of space and also due to the fact that a long article with an overview of PIE argument structure and current research in this area by the author of this chapter will soon appear (Krisch, forthcoming).

#### *Intransitive, transitive, ditransitive and labile verbs*

**Intransitive** verbs, i.e. verbs that do not govern an accusative, do exist in the AIELs and in PIE. However, most instances of these verbs are transitives used “absolutely”, i.e. without an accusative. PIE also has some true intransitive verbs, now termed “unergative”, where the subject has agentive qualities, like the verb ‘to yawn’, PIE \*ǵ<sup>h</sup>i-né/n-h<sub>1</sub>- (LIV<sup>2</sup>: 173). Another group of PIE intransitives are the “unaccusatives”, which have a subject exhibiting the properties of a direct object, e.g. PIE \*key- ‘lie (outstretched)’ (LIV<sup>2</sup>: 20). The verbs of movement are situated between these two classes, since, on the one hand, when moving takes place, there is normally agentive activity involved but, on the other hand, with these verbs the mover and the moved one are identical, e.g. PIE \*h<sub>1</sub>ey- ‘go’ (LIV<sup>2</sup>: 232). Unergatives show a tendency to take active endings (e.g. PIE \*wiwek<sup>w</sup>-ti ‘s/he is speaking’), whereas unaccusatives take stative/middle voice (=medio-passive) endings (e.g. PIE \*key-(t)oy ‘s/he lies’)

**Transitive** verbs, i.e. the verbs that govern an accusative, can be reconstructed for PIE. Also, the “absolute” use of transitives without an overt accusative is a heritage from PIE (see Krisch forthcoming for examples of AIELs with the root PIE \*g<sup>w</sup>h<sup>en</sup>- ‘beat/kill someone’ (LIV<sup>2</sup>: 218)).

**Ditransitive** verbs take a subject, object and “indirect object” and express transfer. Both of the objects either may be expressed or may remain unspecified in the AIELs and in PIE. Examples are PIE \*deh<sub>3</sub>- ‘to give something to someone’ (LIV<sup>2</sup>: 105) and \*b<sup>h</sup>er- ‘to carry/ bring something to someone’ (LIV<sup>2</sup>: 76).

**Labile** verbs, sometimes also called P(atient preserving)-labile verbs, are common in a number of modern IE languages but are rare in AIELs, and until now I have not seen a convincing example of a reconstruction. The object of these verbs can also serve as a subject with the same verb without any formal change on the verb (as in English: *John opens the door* vs. *The door opens*). The reason these verbs are so rare in the AIELs: The early IE languages normally use a specific set of endings (the so-called middle or medio-passive endings) in contrast to active endings to indicate a decrease in valency in the way of the “labile” verbs. If one looks at the diachrony of Modern English, where this phenomenon is most conspicuous, the modern language definitely shows more productivity than the older stages in history (cf. Kulikov 2003: 96 with literature). There are also examples of labile verbs in Vedic Sanskrit (Kulikov 2003: 99; 2014). A good case is *svádate* (3 sg. middle/medio-passive): transitive ‘someone makes sweet something’ (RV 3,54,22) vs. intransitive

‘something is sweet’ (RV. 9,74,9). Investigating Vedic Sanskrit, Kulikov (2003: 103–104; 2014: 1145–1147) regards the intransitive forms of a few labile verbs such as *puṣ* ‘prosper; make prosper’ as older and the transitive constructions as derived from constructions originally containing cognate accusatives. Lavidas (2009) demonstrates that a later development of labile verbs can be proven also for Greek.

### *Subjecthood*

This part of the argument structure has received growing attention in the last years. The normal case form for the subject in the AIELs and in PIE was the nominative. But there does exist so-called non-canonical marking for the underlying subject in a number of languages. This phenomenon (sometimes also labelled “quirky subject” or “oblique subject”) is most prominently present in Icelandic, but there is a growing group of researchers around Jóhanna Barðdal (Bergen) who are collecting and analysing instances of this construction from many IE languages (<http://org.uib.no/iecastp/IECASTP/people.htm>, seen 31 October 2016). In one of the latest publications, one finds the following definition for the oblique subject: “With the terms oblique subject construction and dative subject construction, we . . . refer to constructions where the so-called logical subject is in an oblique case, for instance the dative” (Barðdal & Smitherman 2013: 29). The authors consider the oblique subject construction inherited from PIE. They claim an inherited predicate structure (cognate argument structures; Barðdal & Smitherman 2013: 53) and also present some cognate sets. This seems to me the right way to argue. One of their examples for a cognate set with this kind of construction is the Germanic-Latin-Ancient Greek corresponding construction of the PIE adjective \**sweh<sub>2</sub>d-u-* ‘sweet’:

(60) (cf. Barðdal & Smitherman 2013: 42)

a. Old English, Blick. Hom, 17: 74–75  
 & *him swete wæron to aræfenne*  
 and him.DAT. sweet were to perform.INF.  
 ‘and he enjoyed performing’

b) Latin, Lucr. 4,658  
*quod suaue est aliis*  
 which sweet is others.DAT.  
 which others like/find sweet

c) Ancient Greek (Homer), Od. 24,435–436  
 οὐκ ἄν ἐμοί γε μετὰ φρεσὶν ἦδὺ γένοιτο ζώμεν  
 not PTCL. me.DAT. at.any.rate with souls.DAT. sweet becomes.3SG.OPT. live.INF.  
 ‘Then in my heart, I would not find it at all sweet to live.’

### CONCLUDING REMARKS

After touching upon the history of research and the data situation, this chapter tried to present major fields of study in PIE syntax today: word order; absolute constructions; case functions, including syncretism; and argument structure. Because of the author’s own research interests, the section on word order was especially extensive. The field of PIE syntax radiates in many directions and hopefully will lead to a coherent picture of the syntactic properties of the proto-language.

## FURTHER READING

The text of this chapter has already contained a number of references for the topics dealt with here. One of these references, the English translation of Wackernagel's lectures on syntax (1920, 1924) by Langslow (2009), is an especially interesting read because it combines a masterly translation of one of the most influential books on comparative syntax of the classical languages and Germanic of the 20th century with many footnotes, comparing Wackernagel's ideas with modern viewpoints. Götz Keydana, who has been cited with a number of contributions in this chapter, also wrote an article with a survey of PIE syntax that has been published on the internet, which is worth reading (Keydana 2008).

Some 20 years ago, the introductory textbooks on IE linguistics either did not have chapters on syntax at all (e.g. Szemerényi 1996) or devoted only two or three pages to it (Beekes 1995: 93–95). In more recent introductions, the chapters on syntax are still shorter than those on phonology and morphology, but one can observe an increasing interest in integrating syntax into the picture of PIE. I can recommend Clackson's (2007: 157–186) chapter on PIE syntax, where general methodological questions of syntactic reconstruction, word order (including Wackernagel's Law), clause-linking, alignment change and PIE phraseology are treated. There is also a very useful "Further Reading" section. Fortson (2010: 152–169) deals with topics like constituency (tmesis), subject-verb agreement, Wackernagel's Law, negation, absolute constructions and the interaction of prosody and syntax followed by a short list of "further readings". Meier-Brügger (2003: 238–276; the syntax part is by Matthias Fritz) offers a short discussion of and many bibliographical references to (morpho-)syntactic issues like parataxis and hypotaxis, non-finite constructions, tense-aspect-mood, diathesis, case and gender.

## SOME ABBREVIATIONS USED

Abs.: absolutivum/gerund; AC: absolute construction; AIELs: ancient Indo-European languages; AP: adjectival phrase; C or COMP: complementizer (= subordinate conjunction not forming a phrase in the sentence, e.g. *that, if, because*; this slot may also be occupied by the finite verb, see the German generative tradition; cf., e.g., Haider 2010: 57); CoT: change of topic particle; CP: complementizer phrase; Df-slot: discourse-functional slot; DP: determiner phrase; Mid.: middle (=medio-passive); NP: noun phrase; O: object; t: trace left behind when syntactic movement occurs; PP: prepositional phrase; PRO: empty subject, e.g. in infinitival constructions; pro: "small pro", non-overt noun phrase; RP: relative pronoun; S: subject; SConn: sentence connector; V: verb; V-1: verb in sentence-initial position; V-2: verb in second position; V-3, verb in third position; VP: verb phrase; WL<sub>2</sub>, Wackernagel clitics serving as sentence connectors; WL<sub>1</sub>, other Wackernagel clitics (pronouns); WL<sub>3</sub>, particle adjacent to the word or constituent it takes scope over; XP: any phrase.

## NOTES

- 1 I thank Stefan Niederreiter for proofreading the text. The remaining mistakes are mine, of course.
- 2 Cf., e.g., the various reconstructions of Schleicher's fable in [http://en.wikipedia.org/wiki/Schleicher's\\_fable#L.C3.BChr\\_.282008.29](http://en.wikipedia.org/wiki/Schleicher's_fable#L.C3.BChr_.282008.29) [seen 31 October 2016] or in Birkhan 1985: 307–308. If one compares the syntax of Schleicher's reconstruction from 1868 (cited in Lehmann & Zgusta 1979: 456), *dadarka akvams* (see.3SG.PERF.ACT.

horse.ACC.PL.), and Peters' reconstruction from 1980 (in Birkhan 1985: 308), *dedorke* (*h<sub>1</sub>*)*ekwōns* (see.3SG.PERF.ACT. horses.ACC.PL.), with the reconstruction of Lehmann and Zgusta, *ekwōns espeket* (horses.ACC.PL. catch.sight.of.3SG.AOR., Lehmann & Zgusta 1979: 462), there is a V(erb)-O(bject) order in the first two versions and an OV order in the last one. The OV order clearly reflects Lehmann's syntactic views (cf. also above). One may find a very witty example for a "new" PIE text in the starting episode of the Austrian film *MA 2412. Die Staatsdiener* (MA 2412. The civil servants) by Harald Sicheritz (2003). The PIE text of this film (with German subtitles) was provided by David Stifter (now a professor at Maynooth, Ireland) (2003).

3 For the meaning and origin of the Greek particle τάρ, cf. Katz 2007.

4 A much more complete version of the project can be accessed through the internet (Hettrich 2007). There is much other recent work on functions of cases in AIELs, including monographs (e.g. Hajnal 1995, especially on Myc. *-pi*; Luraghi 2003 on prepositional phrases in Greek) and articles on the accusative (e.g. Boley 2002, Gotō 2002), the dative (e.g. Nowicki 2002, Hettrich 2006) and the genitive (Neumann 2001, Serbat 1992). For Ancient Greek (mainly prepositional phrases) I can recommend a number of very thoroughly argued articles by Conti (e.g. 1996, 2003).

## REFERENCES

- Agbayani, Brian & Golston, Chris 2010, "Second position is first position: Wackernagel's Law and the role of clausal conjunction," *IF* 115: 1–21
- Axel, Katrin 2007, *Studies on Old High German Syntax. Left sentence periphery, verb placement and verb-second*, John Benjamins, Amsterdam–Philadelphia (=Linguistics Today 112)
- Barbosa, Maria do Pilar Pereira 1995, *Null-subjects. PhD Dissertation*, Cambridge, MIT [<http://www.ai.mit.edu/projects/dm/theses/barbosa95.pdf>, seen 31 October 2016]
- Barðdal, Jóhanna & Thomas Smitherman 2013, "The quest for cognates: A reconstruction of oblique subject constructions in proto-Indo-European," *Language Dynamics and Change* 3: 28–67
- Bauer, Anna 2011, "Verberstellung im Hethitischen," in: Krisch, Lindner, Müller et al. (eds.): 39–48
- Beekes, Robert S.P. 1995, *Comparative Indo-European Linguistics. An Introduction*, John Benjamins, Amsterdam–Philadelphia
- Birkhan, Helmut 1985, *Etymologie des Deutschen*, Peter Lang Bern, Frankfurt/Main–New York
- Boley, Jaqueline 2002, "The function of the accusative in Hittite and its implications for PIE," *IF* 107: 124–151
- Caland, Willem 1970 = 1919, *Jaiminīya-Brāhmana*, Teilausgabe übers. von Willem Caland, Sändig, Wiesbaden (=Koninklijke Akademie van Wetenschappen Amsterdam, Afdeling Letterkunde, Nieuwe reeks; 19/4) [reprint from 1919]
- Chantraine, Pierre 1981=1953, *Grammaire Homérique, Tome II Syntaxe. Nouveau tirage*, Éditions Klincksieck, Paris
- Clackson, James 2007, *Indo-European Linguistics. An Introduction*, Cambridge University Press, Cambridge–New York
- Conti Jiménez, Luz 1996, "Zur Verwendung von ἐπί + Dativ und ἐπί + Akkusativ neben Bewegungsverben bei Homer," *MSS* 56: 17–36
- Conti, Luz 2003, "Die Bezeichnung instrumentalischer Bezüge bei Homer," *IF* 108: 195–222



- Croft William 1991, *Syntactic Categories and Grammatical Relations. The Cognitive Organization of Information*, Chicago University Press, Chicago
- Delbrück, Berthold 1967=1893, *Vergleichende Syntax der indogermanischen Sprachen. Erster Theil*, Karl J. Trübner, Strassburg (=Brugmann, Karl & Berthold Delbrück: *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*, 3. Band)
- Delbrück, Berthold 1967=1900, *Vergleichende Syntax der indogermanischen Sprachen. Dritter Theil*, Karl J. Trübner, Strassburg (=Brugmann, Karl & Berthold Delbrück: *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*, 5. Band)
- Dover, Kenneth J. 1960, *Greek Word Order*, Cambridge University Press, Cambridge
- Dressler, Wolfgang Ulrich 1968, *Studien zur verbalen Pluralität: Iterativum, Distributivum, Durativum, Intensivum in der allgemeinen Grammatik, im Lateinischen und Hethitischen*, Böhlau, Wien, Graz (= Sitzungsberichte der Österreichischen Akademie der Wissenschaften, Philosophisch-Historische Klasse 259,1)
- Dressler, Wolfgang 1969, "Eine textsyntaktische Regel der idg. Wortstellung (Zur Anfangsstellung des Prädikatverbums)," *KZ* 83: 1–25
- Dressler, Wolfgang (Ulrich) 1971, "Über die Rekonstruktion der indogermanischen Syntax," *KZ* 85: 5–22
- Fortson, Benjamin W. IV 2010<sup>2</sup>, *Indo-European Language and Culture. An Introduction*, Wiley-Blackwell, Malden–Oxford
- Friedrich, Johannes 1971, *Die hethitischen Gesetze. Transkription, Übersetzung, sprachliche Erläuterungen und vollständiges Wörterverzeichnis. Photomechanischer Nachdruck mit einer Vorbemerkung von A. Kammenhuber*, Brill, Leiden
- Friedrich, Paul 1975, *Proto-Indo-European Syntax: The Order of Meaningful Elements*, Montana College of Mineral Science and Technology, Butte, Mont. (= JIES Monograph Nr. 1)
- Gonda, Jan 1959, "On Amplified Sentences and Similar Structures in the Veda," in: Gonda, Jan (ed.), *Four Studies in the Language of the Veda*, Mouton & Comp, S'Gravenhage (= Disputationes Rheno-Trajectinae III): 7–70
- Gotō, Toshifumi 2002, „Funktionen des Akkusativs und Rektionsarten des Verbums – anhand des Altindoarischen,“ in: Hettrich & Kim (eds.) 2002: 21–42
- Greenberg, Joseph H. 1966, "Some universals of grammar with particular reference to the order of meaningful elements," in: Greenberg, Joseph H. (ed.), *Universals of language*, Cambridge, MIT Press (Mass) 1966<sup>2</sup>: 73–113
- Griffith, Ralph, T.H. 1896, *The hymns of the Rigveda*, 2nd ed. Benares [cited after the electronic version <http://www.sanskritweb.net/rigveda/griffith.pdf>, seen 1 November 2016]
- Haegeman, Liliane 1994<sup>2</sup>, *Introduction to Government and Binding Theory*, Blackwell, Oxford (UK)–Cambridge (USA)
- Haider, Hubert 2010, *The Syntax of German*, Cambridge University Press, Cambridge (= Cambridge Syntax Guides)
- Hajnal, Ivo 1995, *Studien zum mykenischen Kasussystem*, de Gruyter, Berlin (= Untersuchungen zur indogermanischen Sprach- und Kulturwissenschaft, Neue Folge 7)
- Hajnal, Ivo 2008, "Indogermanische Dichtersprache und die Grenzen der Rekonstruktion," in: Huber, Brigitte, Volkart, Marianne & Widmer, Paul (eds.), *Chomolangma, Demawend und Kasbek. Festschrift für R. Bielmeier*, Bd. 2, International Insitutte for Tibetan and Buddhist Studies, Halle (Saale): 457–481
- Hale, Mark Robert 1987, "Notes on Wackernagel's Law in the language of the Rigveda," in: Watkins, Calvert (ed.), *Studies in Memory of Warren Cowgill (1929–1985). Papers*

- from the Fourth East Coast Indo-European Conference, Cornell University, June 6-9, 1985, de Gruyter, Berlin–New York: 38–50
- Hale, Mark 1995, *Wackernagel's Law. Phonology and Syntax in the Rigveda* [unpublished manuscript, August 1995, 302 pages]
- Hale, Mark 1996, "Deriving Wackernagel's Law: Prosodic and syntactic factors determining clitic placement in the language of the Rigveda," in: Halpern, Aaron L. & Zwicky, Arnold M. (eds.), *Approaching Second. Second Position Clitics and Related Phenomena*, CSLI, Stanford, California: 165–197
- Held, Warren H. 1957, *The Hittite Relative Sentence*, Linguistic Society of America, Baltimore (Md) (= Language dissertation 55 = Vol. 33, No. 4 (part 2), suppl.)
- Hettrich, Heinrich 2002, "Das Projekt einer Kasussyntax des Ṛgveda: Der Instrumental," in: Hettrich & Kim (eds.): 43–63
- Hettrich, Heinrich 2006, "Zum Dativ im Vedischen und anderen indogermanischen Sprachen," [http://www.phil.uni-wuerzburg.de/fileadmin/04080400/\_temp\_/Beitrag\_Tagungsakten\_Jena\_2006.pdf, seen 31 October 2016]
- Hettrich, Heinrich 2007, *Materialien zu einer Kasussyntax des Ṛgveda*, Institut für Altertumswissenschaften. Lehrstuhl für Vergleichende Sprachwissenschaft, Würzburg [http://www.phil.uni-wuerzburg.de/fileadmin/04080400/Materialien.pdf, seen 31 October 2016]
- Hettrich, Heinrich & Kim, Jeong-Soo (eds.) 2002, *Indogermanische Syntax. Fragen und Perspektiven*, Dr. Ludwig Reichert Verlag, Wiesbaden
- Hock, Hans Henrich 1996, "Who's on first? Toward a Prosodic Account of P2 Clitics," in: Halpern, Aaron L. & Zwicky, Arnold M. (eds.), *Approaching Second. Second Position Clitics and Related Phenomena*, CSLI, Stanford, California: 199–270
- Holland, Gary B. 1986, "Nominal sentences and the origin of absolute constructions in Indo-European", *KZ* 99: 163–193
- Katz, Joshua 2007, "The epic adventures of an Unknown particle," in: Coulter, George, McCullaugh, Matthew, Nielsen, Benedicte, Ruppel, Antonia & Tribulato, Olga (eds.), *Greek and Latin from an Indo-European Perspective*, Cambridge Philological Society, Cambridge: 65–79
- Keydana, Götz 1997, *Absolute Konstruktionen in altindogermanischen Sprachen*, Vandenhoeck & Ruprecht, Göttingen (=Historische Sprachforschung Ergänzungsheft 40)
- Keydana, Götz 2008, *Indo-European Syntax* [manuscript 21 pages, http://www.keydana.de/deutsch/download.php?id=HSK\_ie\_syntax\_Keydana.pdf, seen 31 October 2016]
- Keydana, Götz 2011, "Wackernagel in the language of the Rigveda. A reassessment," *HS* 124: 106–133
- Keydana, Götz & Luraghi Silvia 2012, "Definite referential null objects in Vedic Sanskrit and Ancient Greek," *Acta Linguistica Hafnensia* 44: 116–128
- Kiparsky, Paul 1995, "Indo-European origins of Germanic syntax," in: Battye, Adrian & Roberts, Ian (eds.), *Clause Structure and Language Change*, Oxford University Press, New York–Oxford: 140–169
- Klein, Jared S. 1991, "Syntactic and discourse correlates of verb-initial sentences in the Rigveda," in: Hock, Hans Henrich (ed.) *Studies in Sanskrit Syntax. A volume in honor of the Centennial of Speiser's Sanskrit Syntax*, Motilal Banarsidass, Delhi: 123–143
- Krisch, Thomas 1984, *Konstruktionsmuster und Bedeutungswandel indogermanischer Verben. Anwendungsversuche von Valenztheorie und Kasusgrammatik auf Diachronie und Rekonstruktion*, Peter Lang, Frankfurt/Main–Bern–New York–Nancy (= Europäische Hochschulschriften Reihe 21, Bd. 35)

- Krisch, Thomas 1986, *Überlegungen zur Herkunft und Entwicklung der irrealen Konditionalsätze des Altgriechischen*, Institut für Sprachwissenschaft, Innsbruck (= Innsbrucker Beiträge zur Sprachwissenschaft. Vorträge und Kleinere Schriften, Band 38)
- Krisch, Thomas 1988, "Zur semantischen Interpretation von absoluten Konstruktionen in altindogermanischen Sprachen (v.a. lateinische und griechische Beispiele)," *Scientia*, Innsbruck (= Scientia Band 10)
- Krisch, Thomas 1990, "Das Wackemagelsche Gesetz aus heutiger Sicht," in: Eichner, Heiner & Rix, Helmut (eds.): *Sprachwissenschaft und Philologie. Jacob Wackernagel und die Indogermanistik heute*, Dr. Ludwig Reichert Verlag, Wiesbaden: 64–81
- Krisch, Thomas 1997, "B. Delbrücks Arbeiten zur Wortstellung aus heutiger Sicht," in: Crespo, Emilio & García Ramón, José Luis (eds.), *Berthold Delbrück y la sintaxis indoeuropea hoy, Actas del Coloquio de la Indogermanische Gesellschaft Madrid, 21-24 de septiembre de 1994*, Dr. Ludwig Reichert, Wiesbaden: 283–309
- Krisch, Thomas 1998, "Zum Hyperbaton in altindogermanischen Sprachen," in: Meid, Wolfgang (ed.), *Sprache und Kultur der Indogermanen. Akten der X. Fachtagung der Indogermanischen Gesellschaft Innsbruck, 22.-28. September 1996*, Institut für Sprachwissenschaft Innsbruck: (= IBS 93): 351–384
- Krisch, Thomas 2001, "'Man kann sich ein Klavier ja auch um den Bauch binden'. Können Theorien der allgemeinen Sprachwissenschaft für die Indogermanistik nützlich sein?" in: Heiner, Eichner; Mumm, Peter-Arnold; Panagl, Oswald & Winkler, Eberhard (eds.), *Fremd und Eigen: Untersuchungen zu Grammatik und Wortschatz des Uralischen und des Indogermanischen, in memoriam Hartmut Katz*, Ed. Praesens, Wien: 155–174
- Krisch, Thomas 2004, "Some aspects of word order and sentence type: From Indo-European to New High German," in: Krisch, Thomas, Lindner, Thomas & Müller, Ulrich et al: 106–129
- Krisch, Thomas 2009, "On the 'Syntax of Silence' in Proto-Indo-European," in: Hinterhölzl, Roland & Petrova, Svetlana (eds.), *New Approaches to Word Order Variation and Change in Germanic*, Mouton de Gruyter, Berlin: 192–220
- Krisch, Thomas 2015, "Zur Erarbeitung einer Datenbank für die Wortstellungs- und Valenzbände im Rahmen der indogermanischen Grammatik", in: Krisch, Thomas & Niederreiter, Stefan (eds.), *unter redaktioneller Mitwirkung von Michael Crombach, Diachronie und Sprachvergleich: Beiträge aus der Arbeitsgruppe "Historisch-vergleichende Sprachwissenschaft" bei der 40. Österreichischen Linguistiktagung 2013 in Salzburg*, Institut für Sprachen und Literaturen, Innsbruck (=IBS 150): 67–77
- Krisch, Thomas (forthcoming), "Argument Structure," in: Garrett, Andrew & Weiss, Michael (eds.), *Handbook of Indo-European Studies*, Oxford University Press, Oxford
- Krisch, Thomas, Lindner, Thomas & Müller, Ulrich [unter redaktioneller Mitwirkung von Michael Crombach, Stefan Niederreiter, Helga Panagl und Ursula Pavičić] (eds.) 2004, *Analecta Homini Universali Dicata. Arbeiten zur Indogermanistik, Linguistik, Philologie, Politik, Musik und Dichtung. Festschrift für Oswald Panagl zum 65. Geburtstag*, Verlag Hans-Dieter Heinz, Akademischer Verlag Stuttgart, Stuttgart (= Stuttgarter Arbeiten zur Germanistik Nr. 421)
- Kulikov, Leonid 2003, "The labile syntactic type in a diachronic perspective: The case of Vedic," *SKY Journal of Linguistics* 16: 93–112
- Lambert, Pierre-Yves 1994, *La langue gauloise : description linguistique, commentaire d'inscriptions choisies*, Éd. Errance, Paris (=Collection des Hespérides : Archéologie, histoire linguistique)

- Langslow, David 2009, *Jacob Wackernagel. Lectures on Syntax with Special Reference to Greek, Latin, and Germanic*, ed. with notes and bibliography, Oxford University Press, Oxford–New York
- Lavidas, Nikolaos 2009, *Transitivity Alternations in Diachrony. Argument Structure and Voice Morphology*, Cambridge Scholars Publishing, Cambridge
- Lehmann, Winfred P. 1974, *Proto-Indo-European syntax*, Univ. of Texas Press, Austin (Tx)
- Lehmann, Winfred P. & Zgusta, Ladislav 1979, “Schleicher’s tale after a century,” in: Brogyani, Bela (ed.), *Studies in Diachronic, Synchronic, and Typological Linguistics. Festschrift for Oswald Szemerényi*, Benjamins, Amsterdam: 455–468
- LIV<sup>2</sup> = Kümmel, Martin and Helmut Rix (eds.) 2001. *Lexikon der indogermanischen Verben*. Zweite, erweiterte und verbesserte Auflage. Dr. Ludwig Reichert Verlag, Wiesbaden
- Lühr, Rosemarie 2009, “P2-Partikeln in indogermanischen Sprachen,” in: Riecken, Elisabeth & Widmer, Paul (eds.), *Pragmatische Kategorien. Form, Funktion und Diachronie. Akten der Arbeitstagung der Indogermanischen Gesellschaft vom 24. Bis 26. September 2007 in Marburg*, Reichert Verlag, Wiesbaden: 173–186
- Lühr, Rosemarie & Zeilfelder, Susanne 2011, “Zur Interdependenz von Diskursrelationen und Konnektoren in indogermanischen Sprachen: Kontrast und Korrektur,” in: Breindl, E.; Ferraresi, G. & Volodina, A. (eds.), *Satzverknüpfungen. Zur Interaktion von Form, Bedeutung und Diskursfunktion*. de Gruyter, Berlin: 107–148
- Luraghi, Silvia 1987, “Patterns of case syncretism in Indo-European languages,” in: Ramat, Anna Giacalone; Carruba, Onofrio & Bernini, Giuliano (eds.), *Papers from the 7th International Conference on Historical Linguistics*, Benjamins, Amsterdam (= Current Issues in Linguistic Theory 48): 355–371
- Luraghi, Silvia 1990, *Old Hittite Sentence Structure*, Routledge, London–New York
- Luraghi, Silvia 1995, “The pragmatics of verb initial sentences in some ancient Indo-European languages,” in: Downing, Pamela & Noonan, Michael (eds.), *Word order in Discourse*, Benjamins, Amsterdam: 354–386
- Luraghi, Silvia 2001, “Syncretism and the classification of semantic roles,” *Sprachtypologie und Universalienforschung (STUF)* 54: 35–51
- Luraghi, Silvia 2003, *On the Meaning of Prepositions and Cases. The expression of semantic roles in Ancient Greek*, John Benjamins, Amsterdam–Philadelphia (= Studies in Linguistics Companion Series Vol. 67)
- McCone, Kim Robert 1979, *Aspects of Indo-European sentence patterns and their role in the constitution of the Old Irish Verbal system*, Oxford Phil. Diss (= British Theses Service D 29191 80LI)
- Meid, Wolfgang 1993, *Die Botorrita-Inschrift. Interpretation eines keltiberischen Sprachdenkmals*, Institut für Sprachwissenschaft, Innsbruck (= IBS 76)
- Meier-Brügger, Michael 2003 [With contributions by Matthias Fritz and Manfred Mayrhofer], *Indo-European Linguistics*, Walter de Gruyter, Berlin–New York [translation of the 2002<sup>8</sup> German edition]
- Melchert, Craig 1977, *Ablative and Instrumental in Hittite*, Harvard University [PhD dissertation]
- Miller, D. Gary 1975, “Indo-European: VSO, SOV, SVO or all three,” *Lingua* 37: 31–52
- Neu, Erich 1980, *Althethitische Ritualtexte in Umschrift*, Harrassowitz, Wiesbaden (=StBot 15)

- Neumann, Günter 2001, "Der adverbale Genitiv im Althethitischen," in: Wilhelm, Gernot (ed.), *Akten des IV. Internationalen Kongresses für Hethitologie Würzburg 4.–8. Oktober 1999*, Harrassowitz, Wiesbaden: 446–455
- Nowicki, Helmut 2002, "Zum Dativ im Althethitischen," in: Hettrich & Kim (eds.): 65–77
- Otten, Heinrich 1973, *Eine althethitische Erzählung um die Stadt Zalpa*, Otto Harrassowitz, Wiesbaden (= Studien zu den Boğazköy-Texten Heft 17)
- Otten, Heinrich & Souček, Vladimír 1969, *Ein althethitisches Ritual für das Königspaar*, Harrassowitz, Wiesbaden (=StBoT 6)
- Pfohl, Gerhard 1980<sup>2</sup>, *Griechische Inschriften als Zeugnisse des des privaten und öffentlichen Lebens*, 2. verb. Aufl., Heimeran, München
- Ruppel, Antonia 2013, *Absolute Constructions in Early Indo-European*, Cambridge University Press, Cambridge (=Cambridge Classical Studies)
- Schmitt, Rüdiger 1967, *Dichtung und Dichersprache in indogermanischer Zeit*, Harrassowitz, Wiesbaden
- Serbat, Guy 1992, "Zum Ursprung des indogermanischen Genitivs und seiner lateinischen Verwendung," in: Panagl, Oswald & Krisch, Thomas (eds.), *Latein und Indogermanisch. Akten des Kolloquiums der Indogermanischen Gesellschaft. Salzburg, 23.–26. September 1986* (= Innsbrucker Beiträge zur Sprachwissenschaft Bd. 64): 285–291
- Sicheritz, Harald 2003, Die Staatsdiener (MA 2412 The civil servants). Film <https://www.youtube.com/watch?v=2i87oG1WQcg> (seen 31 October 2016)
- Sternefeld, Wolfgang 2006, *Syntax. Eine morphologisch motivierte generative Beschreibung des Deutschen*. 2 Vols., Stauffenberg Verlag Brigitte Narr, Tübingen (= Stauffenberg Linguistik Band 31)
- Stifter, David [2003], Mag. David Stifter 'MA 2412 – Die Staatsdiener', Institut für Sprachwissenschaft, Wien, manuscript [dated 23.4.2003]
- Szemerényi, Oswald 1996, *Introduction to Indo-European Linguistics*, Clarendon Press, Oxford
- Wackernagel, Jacob 1892, "Über ein Gesetz der indogermanischen Wortstellung," *IF* 1: 333–436
- Wackernagel, Jacob 1920, *Vorlesungen über Syntax mit besonderer Berücksichtigung von Griechisch, Lateinisch und Deutsch. Erste Reihe*, Verlag Emil Birkhäuser&Cie, Basel
- Wackernagel, Jacob 1924, *Vorlesungen über Syntax mit besonderer Berücksichtigung von Griechisch, Lateinisch und Deutsch. Zweite Reihe*, Verlag Emil Birkhäuser&Cie Basel
- Watkins, Calvert 1963, "Preliminaries to a historical and comparative analysis of the syntax of the Old Irish verb," *Celtica* 6: 1–49
- Watkins, Calvert 1964, "Preliminaries to the reconstruction of Indo-European sentences structure," in: Lunt, Horace (ed.), *Proceedings of the Ninth International Congress of Linguistics Cambridge, Mass., August 27–31, 1962*, Mouton, Cambridge (MA)–London (= Janua linguarum Series maior 12): 1035–1045
- Watkins, Calvert 1976, "Towards Proto-Indo-European syntax: Problems and pseudo-problems," in: *CLS, Papers from the Parasession on Diachronic Syntax*, ed. by Steever, Sanford B., Walker, Carol A., Mufwene, Salikoko S. Chicago Linguistic Society, Chicago: 305–326
- Watkins, Calvert 1995, *How to Kill a Dragon. Aspects of Indo-European Poetics*, Oxford University Press, New York–Oxford

# PROTO-INDO-EUROPEAN AND LANGUAGE TYPOLOGY

*Ranko Matasović*

## INTRODUCTION

This chapter will give an overview of the typologically relevant features of Proto-Indo-European. By “typologically relevant” we mean those features that are cross-linguistically rare, or those that are correlated with other features, thus defining a linguistic “type”, a set of languages sharing a number of logically independent features. Moreover, we will consider those features of PIE that show genetic and/or areal patterning, in the sense that they statistically tend to be more common in certain linguistic families and/or language areas than in others. Particularly relevant are diachronically stable features that tend to be shared by most genetically related languages of a single family. Such features are important for questions about the deep genetic relatedness of Indo-European and about early language contacts between PIE and its neighbours. It has been long suggested that Indo-European languages and Uralic languages are related, forming a putative Indo-Uralic family (e.g. Collinder 1960), and Vladimir Illič-Svityč (1971–84) included both Indo-European and Uralic into a wider family, “Nostratic”, which would also include Altaic, Dravidian, Kartvelian and Afro-Asiatic languages. Several variants of this hypothesis are currently in circulation, including the daring proposal by Joseph Greenberg (2000), who also includes Gilyak, Ainu, Chukchi-Kamchatkan and Eskimo-Aleut (but excludes Afro-Asiatic, Kartvelian and Dravidian) in his “Eurasian” language family. None of these hypotheses can be considered as proven, but many linguists believe that Indo-European is indeed related to Uralic, and that it is more likely to be related to Altaic and Kartvelian than, say, to NW Caucasian or Sino-Tibetan. As this chapter will show, PIE shares a number of typologically relevant features with Uralic and, to a lesser extent, with Altaic and other language families of Northern Eurasia. However, PIE also shares a considerable number of features with languages of the Caucasus, including not only Kartvelian – which is included in the putative Nostratic macrofamily – but also NE and NW Caucasian languages, to which it is probably unrelated. These shared features can be understood only as results of prehistoric language contacts, so they are very relevant for the question of the homeland of the speakers of PIE (Mallory 1989, Anthony 2007).

## PHONOLOGICAL TYPOLOGY

There is far more agreement, among linguists, about the reconstruction of PIE phonology than there is about the reconstruction of the proto-language’s morphology and syntax. Today, it is almost universally agreed that PIE had three series of stops, traditionally labelled as voiceless, voiced and (voiced) aspirated (i.e. \*T, \*D and \*D<sup>h</sup>, respectively). The phonological interpretation of these segments remains an open matter. The traditional

interpretation has been challenged by Roman Jakobson (1962) and other linguists, since there are apparently no languages that have voiced aspirates unless they also have voiceless aspirates. Moreover, the very term “voiced aspirates” is somewhat misleading, since consonants that are represented (and transcribed) as /b<sup>h</sup>, d<sup>h</sup>, g<sup>h</sup>/, etc., in Hindi and other Indo-Aryan languages actually have breathy voice. It is still not clear whether there are languages with voiceless and voiced stops as well as those with breathy voice (but without voiceless aspirated stops), but if such languages exist, they are certainly very rare. Moreover, it has been pointed out that the apparent rarity, or non-existence, of PIE \*b is at odds with the typological generalization that if one of the labial stops is missing in a language with a voice opposition in stops, it is the voiceless labial /p/ rather than its voiced counterpart /b/ that is missing (e.g. in Classical Arabic). These facts led to a phonetic re-interpretation of stops in PIE, known as “glottalic theory”, first proposed by Paul Hopper and, independently, by Tamaz Gamkrelidze and Vyacheslav Ivanov, in the early 1970s (see Gamkrelidze & Ivanov 1984 and Salmons 1994, with references). According to this theory, the stops traditionally interpreted as voiced were actually glottalized, presumably ejectives. It has been noted that the bilabial ejective stop (/pʼ/, corresponding to the \*b of traditional Indo-European reconstruction) is highly marked, and that it is sometimes lacking in languages having ejective consonants. The change of glottalized stops into voiced stops, which has to be posited for several branches of PIE (e.g. for Italic, Celtic and Slavic), is well attested in a number of languages, and the developments in Germanic and Armenian (where voiceless stops correspond to PIE \*b, \*d, \*g, etc., of the traditional reconstruction) are seen as simple deglottalization. Moreover, certain sound developments, such as Winter’s Law in Balto-Slavic (the lengthening of vowels before voiced stops, perhaps only in closed syllables), have been claimed to be more plausible, from the phonetic point of view, if the glottalic theory is correct. However, the glottalic theory has not been generally accepted by Indo-Europeanists, and many scholars prefer to remain agnostic about the correct phonetic interpretation of the PIE system of stops.

In terms of the number of segments, PIE was well within the average of Eurasian languages. However, the number of stops, with three series and three different types of gutturals (plain, palatalized and labiovelar), is unusual, as such complex systems are found mostly in the languages of the Caucasus and in those of South Asia and SE Asia. If one accepts the glottalic theory, it is highly significant that Caucasus is the only area in Eurasia where ejective consonants are common. They even spread to some intrusive languages, such as dialects of Armenian and Ossetic.

One disputed matter in PIE phonology is directly relevant to the typological profile of the proto-language. It is the number of vowels: in traditional handbooks, it is usually assumed that PIE had the typologically most widespread five-vowel system, with a length contrast, i.e. that there were ten vocalic phonemes (\*a, \*e, \*i, \*o, \*u, \*ā, \*ē, \*ī, \*ō, \*ū; cf., e.g., Cowgill & Mayrhofer 1986, Szemerényi 1989). However, the discovery of laryngeals has allowed Indo-Europeanists to show that \*ā, \*ī and \*ū were actually sequences of short \*a, \*i and \*u and laryngeals (i.e. that we should reconstruct \*eh<sub>2</sub> instead of \*ā, and \*iH and \*uH instead of \*ī and \*ū respectively), and it has been suggested that all instances of alleged PIE \*a can be eliminated, either by positing new sound laws or by assuming that etymons with this vowel were early borrowings from unknown sources, chiefly in the languages of the European branches of Indo-European (Lubotsky 1989). It has also been argued that the PIE long vowels \*ē and \*ō were certainly rare, occurring only in a number of morphological categories, such as the sigmatic aorist and the nominative singular of certain nominal stems, and it appears likely that they developed in Late PIE by lengthening of \*e and \*o, respectively, in certain environments. For example, the

long vowel in the nom. sg. of \*ph<sub>2</sub>tēr ‘father’ has been plausibly derived from earlier short \*e in Early PIE \*ph<sub>2</sub>ter-s, with the nom. sg. ending \*-s, by compensatory lengthening, i.e. with the development \*-ers > \*-err > \*-ēr. Furthermore, if the high vowels \*i and \*u are analysed as syllabic allophones of the glides \*y and \*w, we are left – at least in Early PIE – with a binary vocalic system with only \*e and \*o, typologically similar to the binary systems found in NW Caucasian languages such as Abkhaz and Ubykh. Phonetically, \*e and \*o could have been /ə/ and /a/, since languages with two-vowel systems appear always to oppose a low and a mid vowel, rather than a front and a back vowel.

PIE vocalism was characterized by productive ablaut. In PIE, this type of non-linear morphology was used both in the nominal inflexion (to distinguish case forms, e.g. nom. sg. \*pōd-s vs. acc. sg. \*pod-ṃ and gen. sg. \*ped-os or \*pd-os) and in the conjugation of verbs (e.g. to distinguish the singular forms of athematic presents from their plurals, cf. Skr. *ásmi* ‘I am’ < PIE \*h<sub>1</sub>esmi vs. *smás* ‘we are’ < PIE \*h<sub>1</sub>smes). Similar morphologically conditioned vowel alternations exist in several language families of Eurasia, most notably in Afro-Asiatic and Kartvelian (Gamkrelidze & Mačavariani 1982). However, in those language families ablaut is far more productive in the verbal system than in the inflection of nouns.

It is very likely that PIE had a simple, free tone accent, similar to that found in Vedic, where two phonological tones (high and low) must be distinguished. In Ancient Greek, the accent was tonal, but its position was wholly predictable on verbs and partially predictable on nouns. The accent in Latin was described as tonal by Roman grammarians, but its position is completely predictable and it is likely that all Italic languages went through a stage with a dynamic accent which caused syncope and the reduction of vowel timbres in unaccented syllables. The evolution of tonal accent in Balto-Slavic is still controversial: most scholars would posit an opposition between an acute and a circumflex on long vowels and diphthongs, and a simple high tone on short syllables, but it is unclear whether this system developed from PIE tonal oppositions or arose independently after the loss of PIE laryngeals and some Balto-Slavic phonological processes such as Winter’s Law. The nature of accent in Anatolian and Tocharian is unclear, and all the other early Indo-European languages had dynamic accent (stress). With very few exceptions, modern Indo-European languages have dynamic accent; the old tonal oppositions were preserved in parts of the South Slavic area and in the remaining Baltic languages, and new tonal accents arose independently in parts of Scandinavia and in some West Germanic dialects (Franconian). On the whole, languages with tonal accent are rare in Eurasia: Altaic, Uralic, Kartvelian, Afro-Asiatic, Dravidian and most languages of Siberia (with the exception of Ket, which belongs to the small Yenisean family) are characterized by dynamic accent, while complex tonal systems are clearly an areal phenomenon in SE Asia, in languages belonging to different families such as Sino-Tibetan (e.g. Mandarin and Myanmar), Austro-Asiatic (e.g. Vietnamese), Daic (e.g. Thai), etc. If PIE had tones, its accentual system was certainly not similar to systems found in SE Asia, but it may have been similar to the simple tonal system reported for the NW Caucasian language Abaza and some NE Caucasian languages such as Ingush.

## MORPHOLOGICAL TYPE, WORD CLASSES AND LEXICAL TYPOLOGY

PIE was a synthetic, fusional language, which means that its words were polymorphemic and that single morphemes could express more than one grammatical category (e.g. the Lat. dat. pl. ending *-ibus* in *ped-ibus* ‘feet’ cannot be segmented into separate morphemes



for case and number). In this respect, it was similar to Afro-Asiatic and Kartvelian languages, which are also predominantly fusional, while the North Caucasian, Uralic and Altaic languages are mostly agglutinative (e.g. the Hungarian acc. pl. *ember-ek-et* ‘people, men’ contains separate morphemes for plural, *-ek-*, and accusative, *-et*). It has been claimed that there are traces of an early, more agglutinative morphology in PIE (cf., e.g., Szemerényi 1989): for example, the loc. pl. ending of the *o*-stems *\*-isu* has been segmented into the plural morpheme, *\*-i-* found in pronouns, and the locative ending, *\*-su*, found in consonant stems, which developed from an earlier postposition. Likewise, the first person plural ending *\*-mes/-mos* (e.g. Gr. Dor. *-μες*, Lat. *-mus*) may have been originally segmentable into the first person marker *\*-m* and the plural marker *\*-es/-os*, but such analyses are very speculative. In terms of the level of synthesisism, PIE had 4–5 categories per word, which is the most common type cross-linguistically according to WALS (feature 22A, “Inflectional synthesis of the verb”), in between the extreme cases of the radically isolating languages of SE Asia (e.g. Vietnamese, with 0 categories per word) and polysynthetic languages, such as Circassian (NW Caucasian), with as many as 12 categories per word.

Like Uralic, Altaic, Kartvelian and NE Caucasian languages, Indo-European languages are dependent-marking (Nichols 1992), which means that the dependent element of a syntactic constituent is morphologically marked, rather than its head. Dependent-markedness can be realized on different levels:

- 1) in dependent-marking adpositional phrases, the governed noun is case-marked, while the governing adpositions are unmarked, e.g. in the Latin phrase *in urbe* ‘in the city’, where the noun *urbs* ‘city’ is in the ablative singular, and the preposition *in* is unmarked. The converse of this is the head-marked adpositional phrase, where the adposition is marked (usually by person and/or gender markers), while the noun is unmarked (such a construction exists, for example, in the NW Caucasian language Abkhaz).
- 2) in the dependent-marked possessive construction, the possessed noun is marked (usually by the genitive case), while the possessor is unmarked, e.g. in Lat. *domus uirī* ‘the man’s house’. In a head-marking possessive construction, the possessed noun would be unmarked, while the possessor would get a possessive affix, e.g. in Hungarian *az ember ház-a* (the / man / house-his) ‘the man’s house’.
- 3) on the clause level, in dependent-marking languages objects are case-marked (usually by the accusative case), while verbs are unmarked if they have direct objects, e.g. in Lat. *uide-ō uir-um* (see-1sg.pres / man-acc.sg.) ‘I see the man’. In a head-marking construction, the object would not be case-marked, while the verb would express both the person of the subject and the person of the object (again, this is what we find in Abkhaz, among other languages).

On the whole, dependent-marking languages predominate in Eurasia, but there are exceptions like the NW Caucasian languages and the isolates Burushaski (in the Hindukush) and Ket and Nivkh (in Siberia).

Like the vast majority of languages of Northern Eurasia, PIE was a predominantly suffixing language. Whether it had any prefixes at all, apart from the negative prefix *\*h<sub>2</sub>-* (> Lat. *in-*, Gr. *ἀ-*, Skr. *a-*), is questionable. The prefix *\*ni-*, identified in *\*ni-sdo-* ‘nest’ (Lat. *nīdus*, etc., from the root *\*sed-* ‘sit’ > Lat. *sedeō*, OCS *sěsti*, Gr. *ἕζομαι*, etc.) is an often mentioned candidate, but it is possible that it should actually be interpreted as a nominal element *\*h<sub>2</sub>ni-*, a locative singular of the same root found in Gr. *ἄνω* ‘up’, Lith. *nuo*

‘from’ and OCS *na* ‘on’. In that case, we should not reconstruct a prefixed noun *\*ni-sdo-* but rather an old compound *\*h<sub>2</sub>ni-sd-o-* ‘upward seat’ > ‘nest’. There was one infix in Late PIE, the present-stem infix *\*-ne-/-n-*, e.g. in Lat. *tangō* ‘I touch’ vs. the perfect *tetigī* (without the infix), or the Skr. 1 sg. pres. *yunájmi* ‘I harness’ vs. the aor. *áyojam* (without the infix). It is probable that this present-stem infix developed from an earlier nasal suffix, by a metathesis in certain types of roots, cf. Lat. *pandō* ‘I spread’, which developed from *\*pat-n-* by metathesis (Gr. *πίτνημι* < PIE *\*pt-né-h<sub>2</sub>-*). Verbs with a nasal present in PIE almost universally have roots ending in stops, never in resonants or *\*s*, so the metathesis was perhaps regular in structures of the type *\*C<sub>1</sub>VC<sub>2</sub>-n- > \*C<sub>1</sub>VnC<sub>2</sub>-*, where *C<sub>2</sub>* was a stop.

Almost all the Indo-European languages remained predominantly suffixing, with the possible exception of modern Insular Celtic languages such as Irish, which is classified as equally prefixing and suffixing by WALS (feature 26A). However, some prefixes developed in nearly all branches of Indo-European, mostly in the verbal system, e.g. the aspectual prefixes in Slavic, or the Germanic verbal prefixes with directional meaning such as German *bei-*, *auf-*, *ein-*, etc. These prefixes usually developed from adpositions, and this is a process that may be observed in its development in the early history of Greek. In Homer, verbal prefixes are still detachable from the verbal root, and intervening elements occur freely. This is the so-called tmesis, e.g. *Γέλασσε δὲ πᾶσα περὶ χθῶν* ‘the whole earth laughed around’, Il. 19.362 (*Γέλασσε . . . περὶ* instead of *περιγέλασσε*). However, in later Greek, the association of prefixes with verbal stems becomes fixed. Adpositions are prepositions in most Indo-European languages, but in Vedic and in the Old Anatolian languages they are postpositions, so most linguists assume that they functioned as postpositions in PIE as well. It has also been argued that PIE did not have any adpositions, but that they developed from adverbs independently in early Indo-European languages (Comrie 1998). Languages without adpositions are rather rare typologically, but there are such languages, especially among the Aboriginal languages of Australia. However this may be, it is worth noting that many, and perhaps most, PIE adpositions can be derived from earlier nouns, petrified in some grammatical case. For example, the adposition *\*h<sub>2</sub>enti* ‘against, in front of’ (Gr. *ἀντί*, Lat. *ante*) is actually the petrified loc. sg. of a noun preserved in Anatolian, which meant ‘face’ (Hitt. *ḫanza* < PIE *\*h<sub>2</sub>ent-s*).

The lexicon of PIE and all of the early attested Indo-European languages was verb-based, rather than noun-based (Nichols 2007). This means that simplex verbs were an open class and many nouns were derived from verbal roots, e.g. the word for ‘word, utterance’, *\*wōk<sup>w</sup>-s* (Skr. *vāk-*, Lat. *uōx*), is derived from the root *\*wek<sup>w</sup>-* ‘to say, utter’ (LIV 614, cf. Ved. *vívakti* ‘says’); the word for ‘eye’, *\*h<sub>3</sub>ek<sup>w</sup>-* (Lat. *oc-ulus*, OCS *oko*, Gr. Hom. du. *ὄσσε*), is derived from the verbal root meaning ‘to see’ (LIV 264, cf. Gr. fut. *ὄψομαι*, Skr. *īkṣate* ‘perceives’); and the word for ‘wheel’, *\*k<sup>w</sup>ek<sup>w</sup>los* (Gr. *κύκλος*, Skr. *cákra-*), is derived from the root *\*k<sup>w</sup>elh<sub>1</sub>-* ‘to turn’ (LIV 345, cf. Ved. *cáрати* ‘moves’). In noun-based languages, on the other hand, lexical verbs are a small, closed class, and new verbal meanings can be derived by using constructions with auxiliaries (light verbs) and nouns, e.g. in Ingush (NE Caucasian) *Cuo nab ju* (He-erg. / sleep / does) ‘He sleeps’. Of course, PIE also had nouns underived from verbal roots, but they seem to have been a closed class, e.g. *\*k<sup>w</sup>ōn* ‘dog’ (Skr. *śvan-*, OIr. *cú*, Lat. *canis*, etc.), *\*ped-* ‘foot’ (Skr. *pad-*, Lat. *pēs*, etc.). Languages with noun-based lexica include Basque, the Dravidian languages and many NE Caucasian languages, while verb-based lexica are typical of the NW Caucasian languages and (to some extent) Semitic languages.

Cardinal numerals from one to four were inflected for gender, case and number, like nouns and adjectives, and the higher numerals were uninflectable. From the grammatical point of view, the PIE numerals are hardly distinguishable from adjectives; i.e. their

identification as a separate word class is questionable. Likewise, participles were syntactically adjectives derived from verbs. They distinguished different gender forms, just like ordinary adjectives, and were inflected for case and number. The participles included the present active participle (PIE \*b<sup>h</sup>er-o-nt- ‘carrying’ > Skr. *bhárant-*, OCS *bery*, Gr. φέρων, etc.), the perfect participle (PIE \*weyd-wos- > Skr. *vidván-*, Gr. εἰδώς), the present middle participle (PIE \*b<sup>h</sup>er-o-mh<sub>1</sub>nos > Skr. *bhárāmānas*, Gr. φερόμενος) and a verbal adjective in \*-to- with passive meaning (PIE \*b<sup>h</sup>rtós > Skr. *bhṛtá-*, OIr. *-breth*). In areal terms, rich systems of participles are very common in Uralic and Altaic languages, but they are also well attested in many Caucasian languages. Verbal adverbs (also called “gerunds” or “converbs”), forms translatable as ‘while doing X’, where X is the verb from which the adverb is derived, cannot be posited for PIE. They are very common in Altaic and North Caucasian languages.

PIE also had interjections and particles (e.g. the negation \*ne > Lat. *ne*, Skr. *na*, OCS *ne*, etc.), and there may have been some conjunctions, though these are usually clitics in the early Indo-European languages (e.g. Skr. *-ca*, Lat. *-que*, Gr. *-τε* < PIE \*k<sup>w</sup>e ‘and’). These word classes are probably universal, or nearly so. We may note the lack of proper conjunctions in Altaic languages, except for some loanwords from Arabic or Indo-European languages.

Although a lot has been written about the word classes in PIE, there has been relatively little research into the lexical typology of the proto-language. A recent study on the diachronic typology of motion event encoding (Verkerk 2014) found some evidence that PIE tended towards the satellite-framed type of language, which means that motion events were more commonly encoded by verbs and adpositions, as in English *walk across (the river)*, rather than by verbs alone, as in the verb-framed type of language, such as French, where the equivalent construction is *traverser (la rivière)*. The distinction between satellite-framed and verb-framed languages, proposed by Leonard Talmy (e.g. Talmy 2007) is not absolute, of course, and PIE, like most of its descendants, certainly had both verb-framed and satellite-framed constructions. However, some branches tended to develop more towards the verb-framed end of the scale (e.g. Romance and Indo-Iranian), while others evolved more towards the satellite-framed type (e.g. Germanic and Balto-Slavic; see Verkerk 2014).

## GRAMMATICAL CATEGORIES

The PIE nouns, adjectives and pronouns distinguished the grammatical categories of case, number and gender. The case system of PIE is fairly typical of a large number of languages and language families of Eurasia. The case system in PIE, with its eight or nine cases (nominative, accusative, vocative, genitive, dative, ablative, locative and instrumental), is comparable to the one found in several Altaic languages (especially in the Turkic branch), but it is significantly smaller than the exuberant case systems, with many local cases, found in Uralic or NE Caucasian. A typologically unusual feature of the PIE case system is the fact that the nominative case of the masculine and feminine gender is marked (usually with the morpheme \*-s), cf. PIE \*h<sub>1</sub>ekwo-s ‘horse’. In most languages with the nominative-accusative alignment system, the nominative is the unmarked case, while the accusative is usually marked by an affix. In PIE, the accusative of the feminines and masculines (and the nominative-accusative of the thematic neuters) is marked with the suffix \*-m (cf. PIE acc. sg. \*h<sub>1</sub>ekwo-m ‘horse’, nom./acc. sg. n. \*yugo-m ‘yoke’). The nominative-accusative of the athematic neuters was unmarked (cf. PIE \*med<sup>h</sup>u ‘mead’). Another typologically interesting feature of the PIE case system is the widespread syncretism of core and non-core inflectional case-marking, the pattern in which two different

case functions are expressed by a single morpheme (WALS feature 28A). In PIE, this type of syncretism affected both core cases (e.g. the nominative, vocative and accusative of neuters) and non-core cases (e.g. the ablative and genitive singular, except in the thematic inflection), but this pattern is rather rare cross-linguistically, as case syncretism usually affects only core cases. However, the PIE pattern of case syncretism is found in some Uralic languages (e.g. Finnic and Nenets) and in some languages of the Caucasus (e.g. Georgian and Lezgian).

Some PIE nouns belonging to the neuter gender (on which see below) had heteroclitic inflection, which means that they formed their oblique cases from a stem different from the stem of their nominative-accusative singular, e.g. PIE \*yēk<sup>w</sup>r̥t̥ ‘liver’ (nom./acc. sg.) vs. \*yek<sup>w</sup>en-s (gen. sg.). This pattern of inflection is preserved only in a few archaic cases in early Indo-European languages (e.g. Lat. *iecur* ‘liver’, gen. sg. *iocineris*, Skr. *yákr̥t*, gen. sg. *yaknás*), but it is still fairly common in the Anatolian branch, and it was probably more widespread in PIE. This kind of inflectional pattern, where a noun has different stems in grammatical and in oblique cases, is rare in Eurasia, but it is well attested in the NE Caucasian family (Matasović 2012; cf., e.g., Dargi *neš* ‘mother’ with oblique stem *neš-li-*, dat. sg. *neš-li-s*).

The number system, with its opposition of singular, dual and plural, is not so common in Eurasia. In many languages, the number oppositions are not obligatorily expressed; i.e. nouns and verbs are inherently neither singular nor plural (nor dual), although there may be affixes indicating number which may be used when this is appropriate (e.g. in Japanese). In PIE, the expression of number was obligatory for all nominal and verbal forms, and the dual number is probably as old as the singular and the plural, although it appears to have been lost in Anatolian (Fritz 2011). The dual is particularly rare in languages of Northern Eurasia: it is found in some Uralic languages (especially in the Samoyedic branch), but not in Altaic, Sino-Tibetan (though there are exceptions) and Dravidian, and it is generally absent in the languages of the Caucasus. However, early Semitic languages had the dual, but it was gradually lost in most of them.

Late PIE had three genders: masculine, feminine and neuter. It is highly probable (although not generally accepted) that this system developed from an earlier one in which only two genders were distinguished, a common gender (for animates and some inanimates) and a neuter gender (for the remaining inanimates). Such a system could be reflected in the Anatolian languages, which may have separated from the other branches of Indo-European before the feminine gender developed (Matasović 2004). The gender system of PIE is typologically somewhat unusual when compared with the gender systems of other language families of Eurasia, particularly with those found in Dravidian and NE Caucasian (e.g. Chechen). In most of them, purely semantic principles of gender assignment are found; i.e. the gender of a noun is predictable only from its meaning. In PIE, on the other hand, gender can be assigned to only a few nouns on the basis of their meaning only: nouns denoting male persons and higher animals are masculine, while nouns denoting females are feminine. For all the remaining nouns, gender is partly predictable from the declension class that the noun belongs to. For example, nouns derived with the suffix \*-ti- are feminine (e.g. \*m̥tis ‘thought, memory’), while those derived with the suffix \*-r/-n- (the heteroclitics) are neuter (e.g. \*yēk<sup>w</sup>r̥t̥ ‘liver’). Surprisingly, the only language in Eurasia with gender assignment principles and three genders, similar to those found in PIE, is Ket, the only surviving language of a small Yeniseian family in Siberia (Matasović 2004).

Most modern Indo-European languages preserved gender, but the reduction of the three inherited genders to a binary system was common, e.g. in Insular Celtic, in most

Romance languages (the neuter gender is partially preserved in Romanian), in Lithuanian and Latvian, and in many Indo-Aryan languages such as Hindi. However, gender was lost almost without a trace in Armenian and in some Iranian languages such as Farsi.

Most language families of North Eurasia are genderless, and it has been shown that gender is a genetically very stable feature, in the sense that languages belonging to genderless families almost never acquire it, while it is seldom lost in languages (Nichols 1992, Matasović 2004). Even in the Indo-European family, gender was completely lost in only a few languages, such as Armenian and a number of Iranian languages. Thus, Altaic, Uralic, Sino-Tibetan, Daic, and most language isolates of Northern Eurasia are genderless, while Dravidian, Afro-Asiatic and NE Caucasian have gender.

PIE pronouns distinguished case, number and gender, just like the nouns and adjectives, but their inflection patterns were very different. It is not only that they had different case/number endings (e.g. the nom. pl. ending \*-i vs. the ending \*-es found in masculine and feminine nouns), but they were characterized by a high degree of suppletivism. For example, the first person singular pronoun had the stem \*(h<sub>1</sub>)eǵ- in the nominative (cf. Lat. *egō*, Gr. *ἐγώ*), but the stem \*(h<sub>1</sub>)me- in the oblique cases (cf. Lat. acc. *mē(d)*, Gr. *ἐμέ*), and one of the demonstrative pronouns had the stem \*so- in the masculine and feminine nominative singular, but \*to- for the oblique cases and the neuter nominative/accusative singular (cf. the opposition between Skr. nom. sg. m. *sas* ‘this’ and nom./acc. sg. *tad*). Moreover, the personal pronouns had different stems in the singular and plural (1 sg. \*h<sub>1</sub>eǵ- vs. 1 pl. \*we-/\*ne-, 2 sg. \*tuH- vs. 2 pl. \*we-/\*yuH-), which is a common pattern in Caucasian languages and in the Uralic languages spoken in Europe (e.g. Finnish and Hungarian). Cross-linguistically, this pattern of suppletivism is also common in other macroareas (e.g. in Africa and parts of South America), but not in most parts of Eurasia (e.g. in Altaic, Sino-Tibetan and most Uralic languages). It is difficult to ascertain how many distance contrasts were found in PIE demonstratives. Early Indo-European languages usually have three, roughly corresponding to Lat. *hic / iste / ille*, but in some languages a two-way contrast is also found (as in the English opposition between *this* and *that*). Cross-linguistically both patterns are equally common and widespread in most of the world (WALS feature 41A), with the two-way contrast slightly predominating in Eurasia.

It has been argued that PIE pronouns distinguished between inclusive and exclusive forms of the first person plural (e.g. Lehmann 1993, Gamkrelidze & Ivanov 1984). The exclusive form in PIE would have been built with the stem \*ne- (Lat. *nōs*), while the inclusive would be the stem in \*we- (Skr. *vāyam*, Eng. *we*). This would explain not only why there are two different stems for the first person plural pronoun, but also why the stem \*we- is also found in the second person plural in some languages (e.g. Lat. *uōs*): it could be argued that in those languages the inclusive form (originally referring to both the speaker’s and the hearer’s group) narrowed its reference to the hearer’s group only (i.e. to the plurality of the addressees, the second person plural). However, the hypothesis is strictly unprovable, and it is worth recalling that the opposition between exclusive and inclusive pronouns is unattested in the early Indo-European languages (it developed in a few modern Indo-Aryan languages, such as Gujarati and Marathi, presumably under the areal influence of Dravidian and/or Munda languages, where it is sometimes found). Moreover, it is not characteristic of most language families of Central Eurasia, with the single exception of NE Caucasian languages (it is found, for example, in Chechen).

There is no consensus about the grammatical categories reconstructable for the PIE verb. It is clear that verbs were inflected for person/number of the subject, as in most language families of Eurasia. Agreement with the object (found in languages like Basque,

or the NW Caucasian languages such as Abkhaz) cannot be posited for PIE, nor can we assume that gender distinctions were ever made on PIE verbs (as in Afro-Asiatic or NW Caucasian languages).

The moods reconstructable in Late PIE are (besides the unmarked indicative) the imperative (\*b<sup>h</sup>ere ‘carry!’), the subjunctive (\*b<sup>h</sup>er-e-e-si ‘that you carry’), the optative (\*b<sup>h</sup>er-o-ih<sub>1</sub>-s ‘may you carry’) and, perhaps, the desiderative (\*b<sup>h</sup>er-is-ye-si, which is semantically difficult to distinguish from the optative). Note, however, that the Anatolian languages distinguish only between the indicative and the imperative, and it is not entirely certain that they lost the optative and the subjunctive (although many specialists on Anatolian believe they did). Be that as it may, the subjunctive was also lost in some Indo-European languages (e.g. Balto-Slavic), or it was replaced by new subjunctive formations, e.g. in Celtic and Germanic. In Latin it lost its modal function and became the future tense of the verbs of the 3rd and 4th conjugation (e.g. Lat. *leget* ‘will read’). The optative was also lost in many languages, but in Slavic it became the new imperative (e.g. OCS *beri* ‘take!’, with *-i-* developing regularly from the PIE optative marker \*-ih<sub>1</sub>- added to the thematic stem \*-o-). The desiderative exists as such only in Vedic, but it left its traces in the forms of the future tenses in Old Irish, Lithuanian and elsewhere. It is unclear whether prohibition was expressed by using the imperative and the default negative particle, as in OCS *ne beri* ‘do not take!’, or there was a special prohibitive negation used with some form of the verb other than the imperative, as in Gr. μή φέροις ‘do not carry’ (with the prohibitive negation μή and the optative, rather than imperative, form of the verb). The prohibitive negation \*meh<sub>1</sub>- might be a dialectal form with reflexes in Greek, Indo-Iranian, Albanian and Armenian, but it might also be an archaism preserved in those branches only. Languages with special forms of prohibitive negation and non-imperative forms used in prohibitions are more common in Eastern Eurasia than in its western part (though there are exceptions on the extreme western periphery, such as Irish and Welsh; see WALS feature 71A).

On the whole, the PIE system of moods is not typologically unusual, except for one of its features: according to WALS (feature 73A), the morphological optative (i.e. the morphologically expressed mood whose primary function is expression of wishes) is very rare cross-linguistically, except in some linguistic areas. In the whole of Eurasia, the only such area is the Caucasus, where most languages have a morphological optative (Matasović 2012).

The PIE verb was characterized by three different stems: the present stem, the aorist stem and the perfect stem, cf. the following forms of the first person singular of the verb \*leyk<sup>w</sup>- ‘leave’: \*li-ne-k<sup>w</sup>-mi (present, formed with the nasal infix), \*lik<sup>w</sup>-o-m (aorist), \*(le-)loyk<sup>w</sup>-h<sub>2</sub>e (perfect), corresponding to Gr. λυμπάνω (an innovative thematic present formation, besides equally innovative λείπω), ἔλιπον, λέλοιπα, respectively. The imperfect, which is formed from the present stem by adding the aorist endings and the past tense prefix \*e- (the so-called augment) is probably an innovation of a group of Indo-European dialects. Its reflexes are attested in Greek, Indo-Iranian, Armenian and Phrygian, which also share other dialectal features. It is unclear what was the exact nature of the opposition between the verbal forms derived from the present stem and those derived from the aorist stem, but it is more likely to have been aspectual than temporal. Most linguists believe that the forms derived from the present stems were inherently imperfective, while those derived from the aorist stem were perfective, and that this system was best preserved in Greek. It should be noted, however, that no traces of inherited aspectual oppositions exist in the Anatolian languages, which have a purely tense-based opposition between the present and the preterite. Moreover, the position of the perfect stem in the

system remains somewhat unclear. The perfect may have been, originally, a resultative form, expressing the resultant state of a previous action (Comrie 1998), cf. PIE \*woydh<sub>2</sub>e (Skr. *véda*, Lat. *uīdī*), which is the perfect form derived from the root \*weyd- ‘to see’ (cf. Skr. aor. *ávidam* ‘I saw’). The perfect developed into a tense form in several individual languages, e.g. in Latin, where it merged with the aorist into a single category traditionally called “perfect”, and in Germanic, where it is called the “(strong) preterite”. A typologically similar opposition of three aspectual stems is found in Kartvelian languages, which distinguish between imperfect, aorist and present stems. Languages of the Semitic branch of Afro-Asiatic systematically oppose perfective and imperfective verbal forms with different person/number affixes, which is similar to the PIE distinction between primary (present) and secondary (aorist) verbal endings on the one hand, and perfect verbal endings on the other. Uralic and Altaic languages have tense-based rather than aspect-based verbal systems and generally distinguish between past and non-past tenses. Different tenses are usually formed by adding different suffixes to the verbal root, while the person/number suffixes tend to be the same (or very similar) in all tenses.

The future tense cannot be reconstructed for PIE; it is most likely that the future events were referred to either by using the indicative present tense forms of verbs (as, e.g., in Gothic and in Welsh) or by using one of the subjunctive forms (as in Classical Armenian, where the aorist subjunctive is mostly used). In individual Indo-European languages, new future tenses were created, sometimes by re-interpretation of modal forms (e.g. in Latin or Sanskrit), sometimes analytically by using auxiliary verbs such as ‘want’, ‘become’ or ‘have’, e.g. in Slavic (PSl. \*pěti xotjǫ ‘I want to sing’ > Croat. *pjevat ću* ‘I will sing’) and Germanic (Germ. *ich werde singen*, Eng. *I will sing*). Sometimes the auxiliaries became affixes on main verbs, and thus new synthetic forms of the future came into existence, e.g. in Romance (Lat. *cantare habeo* > Ital. *cantarò*, Fr. *chanterai*).

PIE apparently did not have a proper morphological passive. It is possible, of course, that there was some sort of analytical passive construction, but no clear traces of it have been preserved in the early Indo-European languages. The voice system of PIE is characterized by the opposition between active and middle, whereby the middle voice expresses that the subject is affected by the action of the verb, or that it benefits by it; cf., e.g., Gr. act. *λούω τὰς χεῖρας τοῦ παιδός* ‘I wash the hands of the child’ vs. mid. *λούομαι τὰς χεῖρας* ‘I wash my hands’. Besides this primary function, the middle probably also expressed reflexivity (Gr. *λούομαι* ‘I wash myself’) and reciprocal action for some verbs (e.g. Skr. *vádete* ‘they (dual) talk to each other’). The middle is characterized by a special set of personal endings, which seem to be derived from the active endings, e.g. 3 sg. mid. aor. \*-to vs. act. aor. \*-t, 3 sg. mid. pres. \*-tor vs. act. pres. \*-ti. The similarity of the personal endings of the PIE middle to those of the perfect has long been noted, but there is no consensus about the explanation of this similarity. From the typological point of view, the PIE middle is similar to the grammatical category of “version” in Kartvelian languages, cf. Georgian (“subjective version”) *v-i-c’er* ‘I write (for myself)’ vs. (“neutral version”) *v-c’er* ‘I write’.

## CLAUSE ALIGNMENT

All the early Indo-European languages have nominative-accusative clause alignment, with the partial exception of Hittite. Hittite has a split-ergative alignment, since neuter nouns receive the suffix *-anza* when they are the subjects of transitive verbs, but not when they are the subjects of intransitives or direct objects, cf. Hitt. (KUB XIV (23): *nu KUR<sup>URU</sup> Ḫatti-ya apaš išhan-anza arḫa namma zinnešta* ‘and that murder moreover ended

the land of Hatti'. The common gender nouns in Hittite have the nominative-accusative alignment, with direct objects in the accusative case and all the subjects (of both transitive and intransitive verbs) in the nominative. A similar split-ergative system is attested in the Australian language Mangarrayi, where masculine and feminine nouns have the nominative-accusative alignment, while the neuters have the ergative alignment (Dixon 2002: 508). However, most specialists agree that the ergativity in Hittite is unlikely to be inherited from PIE and that it is probably a dialectal development, perhaps under the areal influence of ergative languages such as Hattic and Hurrian.

Among the modern Indo-European languages, ergative patterns developed in the Iranian languages of the Pamir. Split-ergative systems also developed in some Indo-Aryan languages, such as Hindi, where clauses with perfective verbs have ergative alignment, while those with imperfective verbs have nominative-accusative alignment.

Since early Indo-European languages are overwhelmingly nominative-accusative, it would be natural enough to assume, as many linguists did, that the same situation obtained in PIE. However, as early as 1901, Christian Uhlenbeck proposed to reconstruct the ergative alignment for the proto-language, and he was followed by a substantial minority of Indo-Europeanists, such as Vaillant (1936), Beekes (1995) and Kortlandt (2009), who adduced some new arguments for this hypothesis.

The case for the ergative hypothesis is usually made by pointing out that it allows us to find a single explanation for a number of seemingly unconnected, and typologically somewhat unusual, facts about PIE morphology:

1. That PIE has – at least for the masculine and feminine nouns – a special marker for the nominative case (the nominative is usually the unmarked case in nominative-accusative languages, as we saw above).
2. That there is an unusual system of gender assignment, with many inanimate nouns belonging to the common (masculine or feminine) gender.
3. That the ending of the nominative singular is similar to the ending of the genitive singular.
4. That personal pronouns have very different case endings than nouns.
5. That the stem of one demonstrative pronoun begins with \*s- in the nominative singular of the common gender, but with \*t- in the oblique cases, as well as in all the cases of the neuter gender.

The fact that only nouns denoting animates have the special nominative ending, while the nominative-accusative form of neuters is unmarked, may be explained if one assumes that the inanimates were seldom – if ever – subjects of transitive clauses in Early PIE. Hence, their zero-marked form is derived from the original absolutive, while the nominative form of the animates is the original ergative. Those nouns denoting inanimates that belong to the common gender (e.g. the masculine noun \*pōds 'foot') might have the nominative ending because they generalized the ergative, rather than the absolutive, ending when the clause alignment pattern changed to the nominative-accusative system. Moreover, the nominative ending \*-s of the masculines and feminines is similar to the genitive ending of athematic stems \*-es/-os, which appears to be its ablaut variant. The syncretism of the genitive case with the ergative is well attested in ergative languages, e.g. in Yup'ik (where the genitive is traditionally called *relative*) and in the NE Caucasian language Lak. The reason why there is a "leak" of inanimates into the "animate" gender (i.e. the common gender comprising the masculine and the feminine) might be that the masculine and feminine inanimates are precisely those that generalized their old ergative ending,



rather than their absolute ending, as the new nominative. It has been established that the pattern in which pronouns have nominative-accusative alignment, while the nouns follow the ergative pattern, is quite common cross-linguistically, occurring, e.g., in several Aboriginal languages of Australia, including the well-described Pama-Nyungan language Dyirbal. Finally, the curious heteroclisis of the demonstrative stems may be explained if \*so- is the original ergative case form, while \*to- is the absolute and oblique form. Since the neuters were seldom, if ever, in the ergative case, they generalized the absolute stem \*to- in all their case forms. Of course, all of these arguments in favour of the ergative hypothesis are very speculative, and there may be alternative explanations of the features of PIE morphology that it seeks to explain. Some linguists would even argue that there is nothing to explain, as the development of PIE morphology was subject to many complex factors, including phonological changes that we cannot reconstruct any more.

It has also been argued that clause alignment in Early PIE was active (Lehmann 1974, 1993, Gamkrelidze & Ivanov 1984, Drinka 1999, Bauer 2000). Languages with active clause alignment have two types of intransitive verbs: those expressing actions, such as ‘run’ or ‘jump’, treat their single arguments morphosyntactically as the subject of transitives (if the language has case, the only argument of such verbs would be in the nominative). On the other hand, the single arguments of stative intransitives (e.g. verbs like ‘sit’ or ‘sleep’) are treated as the objects of transitives (if the language has case, the single argument of such verbs would be in the accusative). It has been proposed that traces of active morphosyntax can be found in constructions with “accusative subjects”, which are found in some early Indo-European languages such as Latin, where the verb *paenitet* ‘be discontent with’ takes the accusative (*quemque* ‘each man’):

*Suae quemque fortunae paenitet* (Cicero, *Ad familiares*, 6.1.1)  
‘Each man is discontent with his lot.’

However, as has been argued elsewhere (Matasović 2013b), constructions with such “quirky case-marking” can arise in nominative-accusative languages, and there are clear indications that such constructions in early Indo-European languages are innovations rather than archaisms.

Ultimately, the question of clause alignment in Early PIE probably cannot be resolved; although the evidence that Early PIE had active clause alignment does not carry much weight, the arguments for ergativity are inconclusive, and the hypothesis that it was nominative-accusative, just as Late PIE was, is quite tenable (Villar 1983, Rumsey 1987). From the areal point of view, languages with active clause alignment are quite rare in Eurasia. One such language is Batsbi (also called Tsova-Tush), a NE Caucasian language spoken in Georgia, and Georgian also shows some features of active clause alignment. Nominative-accusative languages predominate in the North and Northeast of Eurasia (notable exceptions are Chukchi in Eastern Siberia, and the isolate Burushaski in Pakistan). Ergative languages are found in the Caucasus (both NE and NW Caucasian languages are overwhelmingly ergative), and most languages of the Ancient Near East, including Hattic, Hurrian and Sumerian, were also ergative, as is Basque, spoken on the western periphery of Europe. The Afro-Asiatic, Uralic and Altaic families are consistently nominative-accusative.

## WORD ORDER, AGREEMENT AND COMPLEX CONSTRUCTIONS

The earliest attested Indo-European languages, Hittite and Vedic Sanskrit, have the basic SOV order, and several linguists have concluded that the same order of syntactic

elements must have been regular in PIE (e.g. Lehmann 1974, 1993, Watkins 1976), cf. Skr. *viśaḥ kṣatṛīyāya balīm haranti* (villagers / to.prince / tax / they.pay) ‘The villagers pay taxes to the prince’. However, others have pointed out that many early Indo-European languages have rather free word order (Friedrich 1975) and that in some of them, e.g. Ancient Greek, the SVO order actually predominates, at least statistically. The Insular Celtic languages are VSO from their earliest attestations, early Germanic languages and Latin are SOV, and Baltic and Slavic languages are mostly SVO.

The basic order of subject, object and verb in PIE has been the centre of the controversy since many linguists believe that the relative order of other elements, such as nouns (N) and adjectives (Adj), demonstratives (Dem) and nouns, relative clauses (Rel) and head nouns, and possessors and possessums (usually represented as nouns and adposed genitives, Gen), can be predicted from it, in accordance with implicational universals posited by Joseph Greenberg (1963, see also Dryer 1992). Thus, it has been argued that in a SOV language we expect to find the orders Dem-N, Adj-N, Rel-N and Gen-N, and postpositions rather than prepositions, as we do in consistently head-final languages such as Turkish and Japanese. If PIE had all the other orders of syntactic elements harmonic with the SOV order, as claimed by Lehmann (1974) and others, including Gamkrelidze and Ivanov (1984), then the inherited construction is found in Hitt. *É-ri anda* ‘in the house’ (with a postposition) rather than in Lat. *in casā* (with a preposition). Moreover, we would expect that the PIE poetic formula ‘daughter of the sky’ is preserved in Ved. *divó duhitā* (with the Gen-N order) rather than in Gr. *θυγάτηρ Διός* (with the N-Gen order), and the Vedic phrase *ákṣiti śrávas* ‘imperishable fame’ (Adj-N) would be more archaic than the synonymous Gr. *κλέος ἄφθιτον* (N-Adj).

However, subsequent work on word order regularities and universals (especially Dryer 1992 and his contributions to WALS) has shown that some alleged word order universals cannot be confirmed if large enough samples of languages are used, and that some word order patterns accidentally co-occur in particular macroareas, but that there is no global implication relating them. For example, the order Dem-N is extremely common in Eurasia (with only marginal exceptions, such as Basque and Insular Celtic, which have the opposite order), so it appears to be correlated with the order OV, which is also very common in Eurasia. However, in other macroareas, the order OV co-occurs with the order N-Dem, and overall there is no implicational universal  $OV \rightarrow Dem-N$ . Therefore, if we assume that objects preceded verbs in PIE, it does not follow that demonstratives must have preceded nouns in the NP. Similarly, adjectives tend to precede nouns in Eurasia, but the order Adj-N is not correlated with the OV order globally, so we cannot infer it from the hypothesis that PIE was an OV language.

Moreover, it is not entirely clear that we are right in attributing the SOV order to PIE in the first place. The order of syntactic elements can change very quickly in the history of a language, it is far less stable than clause alignment or agreement patterns, and it is strongly subject to substratum influences (Thomason 2001). Thus, it is entirely possible that Indo-Iranian acquired the SOV order through contact with some language of the SOV type (both Dravidian and Elamite are basically SOV), and the same goes for the Old Anatolian languages (Hurrian and Akkadian are both quite consistently SOV languages).

No single infinitive form can be reconstructed for PIE (Disterheft 1980), although Indo-European languages usually have it (with the exception of Insular Celtic). Infinitives in early Indo-European languages are clearly derived from petrified case-marked forms of abstract verbal nouns. For example, Latin infinitives in *-ere*, *-āre*, *-ire* are originally locative singular forms of *s*-stems (with rhotacism, the change of  $*s > r$  between vowels,

which is regular in Latin). In Classical Sanskrit, the infinitive in *-tum* (e.g. *bhartum* ‘to carry’) is the accusative singular of the verbal noun derived with the suffix *\*-tu-* (identical to the supine in Slavic, cf. OCS *bratъ* vs. the infinitive *brati* ‘carry’). There are also no subordinating conjunctions (subordinators) that can be posited for PIE, so apparently the proto-language did not have constructions of the English type *I see that he is coming* (with the subordinator *that*) or *I want him to come* (with the infinitive *to come*). Rather, it probably used a number of subordinating strategies involving participles (the English type *I see him coming*) or paratactic instead of hypotactic constructions (*I see it, he is coming*). Many early Indo-European languages have the so-called absolute constructions with participles, which are constructed with one of their arguments in a specific case, e.g. the ablative absolute construction in Latin, cf. *Sōle ortō Uolscī sē corcumuāllātōs uidērunt* (Livius, IV.9.13) ‘when the sun was risen, the Volscians saw that they were surrounded by lines of intrenchment’ (with the expression *sōle ortō* in the ablative absolute). However, the cases used in the absolute constructions do not agree: we find the genitive absolute in Greek and the dative absolute in Old Church Slavic, cf. Gr. τούτων λεχθέντων ἀνέστεισαν ‘this said, they rose’ (Xenophon, Anabasis 3.3.1, with τούτων λεχθέντων in the genitive plural), OCS *narodomъ že sьbirajštemъ se načerъ glagolati* ‘when the people assembled, he started to speak’ (Luke, 11.29, with *narodomъ . . . sьbirajštemъ* in the dative plural). Thus, it is unclear whether a single absolute construction can be posited in PIE.

The use of infinite verbal forms, rather than finite subordinate clauses, is actually the more common strategy of clause combining in Eurasia. Uralic and Altaic languages (except those strongly influenced by Indo-European) usually do not have subordinators and finite subordinate clauses, and the same holds for both North Caucasian families. The use of finite subordinate clauses is more typical of various branches of Afro-Asiatic, especially the Semitic languages.

Two different forms of relative pronouns, *\*(H)yo-* (> Skr. *yas, yā, yad*) and *\*kʷo-* (> Lat. *quī, quae, quod*), can be reconstructed, and they either represent dialectal forms or were used in different types of relative clauses (Hajnal 1997). One type of relative clause that is well attested in the early Indo-European languages, and that certainly goes back to a PIE prototype, is the so-called correlative construction, equivalent to the English *He/the man who sees the woman, he is a warrior* (rather than *He/the man who sees the woman is a warrior*). This is the default relative construction in Hittite and Vedic, and it can be safely posited for PIE, cf. Hitt. *nu kaš kuit memai n-at zik šakti* (ptcl. / this.one / what / says / ptcl.-it / you / know) ‘You will know what this one says’. Overall, inflected relative pronouns as a device for introducing relative clauses are typologically rather rare. Most language families of Eurasia use uninflectable relative clause markers (e.g. Hebrew *ʔašer* ‘such that’) or participles rather than finite relative clauses (e.g. the North Caucasian languages and most Altaic languages). However, inflectable relative pronouns are found in Kartvelian, cf. Georgian *vinc* ‘he who’.

PIE had extensive agreement on both the clause and NP levels. On the clause level, verbs agreed with their subjects in person and number. An exception to this is the so-called τὰ ζῷα τρέχει construction, attested in Greek and Indo-Iranian: the neuter plural (τὰ ζῷα ‘animals’) triggers singular agreement on verbs (τρέχει ‘runs’ is the third person singular present). This construction is generally believed to be archaic, and it shows that neuter plurals behaved as collective nouns in PIE, i.e. that they were singular for purposes of agreement. On the level of the NP, adjectives, pronouns and some numerals agreed with the head nouns in case, gender and number, cf. Lat. *hic bonus puer* ‘this (m.) good (m.) boy’ vs. *haec bona puella* ‘this (f.) good (f.) girl’. This type of adnominal agreement is

typologically rather rarer than clause-level agreement. Areally, it is more common in South and Southwestern Eurasia, typically in such language families as Afro-Asiatic, Dravidian and NE Caucasian. The Altaic, Uralic and NW Caucasian families typically lack adnominal agreement, except in those languages that underwent extensive influences from languages with this agreement pattern (for example Finnish, where case agreement is usually considered to be the result of the areal influence of Indo-European languages). Although case and number, as grammatical categories, appear to be old in PIE, agreement in case and number (and perhaps in gender) may be of more recent origin. There are archaic traces of uninflectable adjectives in PIE (e.g. the prefixed forms \**h<sub>1</sub>su-* ‘good’ and \**dus-* ‘bad’), and demonstrative pronouns that have different case/number/gender forms seem to be derived from uninflectable deictic particles. For example, the pronoun \**so-* / \**seh<sub>2</sub>-* / \**to-* (Skr. *sas*, *sā*, *tad*) seems to be etymologically related to Hittite uninflectable sentence particles *-šu* and *-ta* (Matasović 2013a).

As Meillet (1937) noted, the difference between transitive and intransitive verbs in PIE appears to have been slight, as most verbs taking two arguments (bi-valent verbs) could be used with one argument only; e.g. verbs such as ‘to love’ and ‘to hate’ could be used without the undergoer argument, as in Catullus’ verse *Odi et amo* ‘I love and I hate’. The government relation between the verb and its arguments was probably loose and subject to semantic, more than strictly syntactic, principles, as verbs could govern their arguments in different case forms with slightly different meanings. For example, the Greek verb κλύω ‘to hear’ (< PIE \**klew-*) could govern the accusative (Homer’s ἐκλυον αὐδὴν ‘They heard a voice’), the genitive (ἐκλυον αὐτοῦ ‘They heard (it) from him’) and the dative (εὐχόμενοι μοι ἐκλυον ‘They obeyed my prayer’, lit. ‘they obeyed to me praying’). No valence-decreasing morphosyntactic operation, such as passive, can be reconstructed in PIE. On the other hand, there is a causative construction, expressed by a verb form marked by the *o*-grade of the root and the suffix \*-eye-, which is a valence-increasing operation, cf. PIE \**mon-eye-* > ‘cause to remember, admonish’ > (Lat. *moneō*) from the root \**men-* ‘remember’ (cf. Lat. *meminī* ‘I remember’, which is an old perfect).

## CONCLUSION

In a general survey of the areal typology of PIE, Johanna Nichols (2007) found evidence that PIE shared more areally diagnostic features with languages of Northern Eurasia than with languages of Southern Eurasia. These features would include the nominative-accusative clause alignment, the lack of polypersonal verbs, the suffixes as person markers on verbs and the low index of verbal inflectional synthesis. However, she acknowledged the presence of some features linking PIE with languages of South Eurasia, including three series of stops and the presence of gender. A later study (Matasović 2012) argued that PIE shared some typologically rare features with languages of the Caucasus, including the presence of a morphological optative and a heteroclitic inflection of nouns, while pointing out that some features of PIE are rather rare in Northern Eurasia (e.g. nominal agreement and the high consonant-to-vowel ratio). Several features that could potentially link PIE with the Caucasus as a linguistic area are disputed, in that they are not generally reconstructed in the proto-language, e.g. the glottalized stops, the ergative clause alignment and the opposition of exclusive and inclusive pronouns. Hence, it is difficult to reach firm conclusions about areal-typological affinities of PIE, and it is probable that no consensus about these matters will be attained until substantial progress is made in the reconstruction of the proto-language.

## FURTHER READING

There are several good introductions to linguistic typology and language universals, e.g. Comrie 1989, Whaley 1997, Croft 2003 and Velupillai 2012, and one can also learn a lot about typology and the areal distribution of typologically relevant features by reading the chapters on linguistic features in WALS. The typological profile of PIE is given in an excellent overview by Bernard Comrie (1998). A (rather too radical) critique of the application of typological methods to PIE reconstruction can be found in Haider 1985. The problem of areal-typological affinities of PIE is discussed in Nichols 2007 and Matasović 2012. On the possible deep genetic relations of PIE the literature in English is scant; besides the rather speculative monograph by Greenberg (2000), there is a joint volume edited by Salmons and Joseph (1998), which offers an overview of the Nostratic hypothesis.

Many controversial features of the reconstructed PIE are treated from the areal and typological points of view in the large monograph by Gamkrelidze and Ivanov (1984). Typological arguments relevant for the phonetic interpretation of the PIE phonological system are discussed in the papers assembled in Vennemann 1989 and in the monograph by Joe Salmons (1994). The syntax of PIE is treated from the typological point of view by Lehmann (1974, 1993), but most of the conclusions of both books are highly controversial. Several articles about the typology of certain PIE syntactic constructions are contained in the volume edited by Paolo Ramat (1980). The category of gender in PIE is typologically and areally compared with other gender systems in Eurasia by Matasović (2004), and the typology of the verbal categories of tense and aspect in PIE and its descendants is the subject of Hewson and Bubenik 1997.

## REFERENCES

- Anthony, David 2007, *The Horse, the Wheel, and Language*, Johns Hopkins University Press, Baltimore
- Bauer, Brigitte M.L. 2000, *Archaic Syntax in Indo-European*, Mouton De Gruyter, Berlin
- Beekes, Robert S.P. 1995, *Comparative Indo-European Linguistics*, Benjamins, Amsterdam
- Collinder, Björn 1960, *Comparative Grammar of the Uralic Languages*, Almqvist & Wiksell, Stockholm
- Comrie, Bernard 1989, *Language Universals and Linguistic Typology*, Blackwell, Oxford
- Comrie, Bernard 1998, "The Indo-European language family: Genetic and typological perspectives," in: Ramat, Anna Giacalone & Ramat, Paolo (eds.), *The Indo-European Languages*, Routledge, London: 74–97
- Cowgill, Warren & Mayrhofer, Manfred 1986, *Indogermanische Grammatik. I. Einleitung, Lautlehre*, Winter, Heidelberg
- Croft, William 2003, *Typology and Universals*, Cambridge University Press, Cambridge
- Disterheft, Dorothy 1980, *The Syntactic Development of the Infinitive in Indo-European, Slavica*, Columbus, Oh
- Dixon, Robert M.W. 2002, *Australian Languages*, Cambridge University Press, Cambridge
- Drinka, Bridget 1999, "Alignment in early proto-Indo-European," in: Justus, C.F. & Polomé, E.C. (eds.), *Language Change and Typological Variation: In Honor of Winfried P. Lehmann on the Occasion of his 83rd Birthday*, Vol II, Institute for the Study of Man, Washington, DC: 464–500

- Dryer, Matthew S. 1992, "The greenbergian word order correlations," *Language* 68: 81–138
- Friedrich, Paul 1975, *Proto-Indo-European Syntax: The Order of Meaningful Elements*, The Journal of Indo-European Studies, Butte (Montana)
- Fritz, Matthias 2011, *Der Dual im Indogermanischen*, Winter, Heidelberg
- Gamkrelidze, Tamaz V. & Ivanov, Vjačeslav V. 1984, *Indoevropskij jazyk i indoevropejcy, I-II*, Izdatel'stvo tbilisskogo universiteta, Tbilissi [English translation: *Indo-European and the Indo-Europeans. A Reconstruction and Historical Analysis of a Proto-Language and a Proto-Culture*, De Gruyter, Berlin 1995]
- Gamkrelidze, Tamaz, V. & Mačavariani, Givi I. 1982, *Sonantensystem und Ablaut in den Kartwelsprachen: Eine Typologie der Struktur des Gemeinkartwelischen*, Günter Narr, Tübingen
- Greenberg, Joseph H. 1963, "Some universals of grammar with particular reference to the order of meaningful elements," in: Greenberg, Joseph H. (ed.), *Universals of Language*, Cambridge, The MIT Press, Massachusetts: 73–113
- Greenberg, Joseph H. 2000, *Indo-European and its Closest Relatives: The Eurasianic Language Family*, Stanford University Press, Stanford
- Haider, Hubert 1985, "The fallacy of typology: Remarks on the PIE stop system," *Lingua* 65: 1–27
- Hajnal, Ivo 1997, "Definite nominale determination im indogermanischen," *Indogermanische Forschungen* 102: 213–239
- Hewson, John & Bubeník, Vit 1997, *Tense and Aspect in Indo-European Languages. Theory, Typology, Diachrony*, John Benjamins, Amsterdam
- Illič-Svityč, Vladimir M. 1971–1984, *Opyt sravnenija nostratičeskix jazykov*, Nauka, Moscow
- Jakobson, Roman 1962, "Typological studies and their contribution to historical comparative linguistics," in: Jakobson, Roman, *Selected Writings*, Vol. 1, Mouton, The Hague: 278–292
- Kortlandt, Frederik 2009, "C.C. Uhlenbeck on indo-European, Uralic and caucasian," *Historische Sprachforschung* 122: 39–47
- Lehmann, Winfred, P. 1974, *Proto-Indo-European Syntax*, The University of Texas Press, Austin
- Lehmann, Winfred, P. 1993, *Theoretical Bases of Indo-European Linguistics*, Routledge, London
- LIV = *Lexikon der Indogermanischen Verben*, Rix, Helmut (ed.), Dr. Ludwig Reichert Verlag, Wiesbaden 2001<sup>2</sup>
- Lubotsky, Alexander 1989, "Against a PIE phoneme \*a," in: Vennemann, Theo (ed.), *New Sound of Indo-European*, De Gruyter, Berlin: 53–66
- Mallory, James 1989, *In Search of the Indo-Europeans*, Thames and Hudson, London
- Matasović, Ranko 2004, *Gender in Indo-European*, Winter, Heidelberg
- Matasović, Ranko 2012, "Areal typology of PIE: The case for caucasian connections," *Transactions of the Philological Society* 110: 283–310
- Matasović, Ranko 2013a, "Nominal agreement in PIE from the areal and typological point of view," in: Neri, Sergio & Schuhmann, Roland (eds.), *Studies on the Collective and Feminine in Indo-European from a Diachronic and Typological Perspective*, Brill, Leiden: 133–155
- Matasović, Ranko 2013b, "Latin paenitet me, miseret me, pudet me and active clause alignment in PIE," *Indogermanische Forschungen* 118: 93–110

- Meillet, Antoine 1937, *Introduction à l'étude Comparative des langues indo-européennes*, Champion, Paris
- Nichols, Johanna 1992, *Linguistic Diversity in Space and Time*, The University of Chicago Press, Chicago
- Nichols, Johanna 2007, "A typological geography for proto-Indo-European," in: Jones-Bley, Karlene, Huld, Martin E., Della Volpe, Angela and Robbins Dexter, Miriam (eds.), *Proceedings of the 18th UCLA Indo-European Conference*, Institute for the Study of Man, Washington, DC: 191–211
- Ramat, Paolo (ed.) 1980, *Linguistic Reconstruction and Indo-European Syntax*, John Benjamins, Amsterdam
- Rumsey, Alan 1987, "The chimera of proto-Indo-European ergativity," *Lingua* 71: 297–318
- Salmons, Joseph 1994, *The Glottalic Theory: Survey and Synthesis*, The Journal of Indo-European Studies Monographs, Institute for the Study of Man, Washington, DC
- Salmons, Joseph & Joseph, Brian D. (eds.) 1998, *Nostratic. Sifting the Evidence*, Benjamins, Amsterdam
- Szemerényi, Oswald 1989, *Einführung in die vergleichende Sprachwissenschaft*, Wissenschaftliche Buchgesellschaft, Darmstadt [English translation: *Introduction to Indo-European Linguistics*, Oxford University Press, Oxford 1996]
- Talmy, Leonard 2007, "Lexical typologies," in: Shopen, Timothy (ed.), *Language Typology and Syntactic Description: Grammatical Categories and the Lexicon*, Cambridge University Press, Cambridge: 66–168
- Thomason, Sarah G. 2001, *Language Contact*, Edinburgh University Press, Edinburgh
- Uhlenbeck, Christian C. 1901, "Agens und patiens im Kasussystem der indogermanischen Sprachen," *Indogermanische Forschungen* 12: 170–172
- Vaillant, André 1936, "L'ergatif indo-européen," *Bulletin de la Société de Linguistique de Paris* 37: 93–108
- Velupillai, Viveka 2012, *An Introduction to Linguistic Typology*, Benjamins, Amsterdam
- Vennemann, Theo (ed.) 1989, *The New Sound of Indo-European, Essays in Phonological Reconstruction*, De Gruyter, Berlin
- Verkerk, Annemarie 2014, *The Evolutionary Dynamics of Motion Event Encoding*, Max Planck Institute for Psycholinguistics, Nijmegen
- Villar, Francisco 1983, *Ergatividad, Accusatividad y género en la familia lingüística indoeuropea*, Universidad de Salamanca, Salamanca
- WALS 2005, *World Atlas of Language Structures*, Haspelmath, Martin, Dryer, Matthew S., Gil, David and Comrie, Bernard (eds.), Oxford University Press, Oxford
- Watkins, Calvert 1976, "Towards proto-indo-european syntax: Problems and pseudo-problems," in: Steever, Sanford B., Walker, Carol A. and Mufwene, Salikoko S. (eds.), *Papers from the Parasession on Diachronic Syntax*, Chicago Linguistic Society, Chicago: 304–317
- Whaley, Lindsay J. 1997, *Introduction to Typology*, Sage Publications, Thousand Oaks

# ANATOLIAN

*H. Craig Melchert*

## INTRODUCTION

### The Anatolian subgroup as a whole

The Anatolian subgroup of Indo-European is attested from the middle of the second millennium BCE to approximately the first and second centuries CE. There are no modern representatives of the family. The evidence ranges from the extensive records for Hittite through the significant but progressively smaller text corpora in Luwian, Lycian, Lydian, Carian, and Palaic to the mere handful of inscriptions in Sidetic and Pisidian. We have Hittite, Palaic, and Luwian texts from the second millennium in the cuneiform writing system borrowed from Mesopotamia and Luwian inscriptions in a native Anatolian system of hieroglyphs from the late second and early first millennium BCE. After a significant break, the remaining languages from the late first millennium are written in alphabets closely related to or derived from that of Greek. The great disparity in the amount of evidence for the respective languages and the nature of the cuneiform and hieroglyphic writing systems limit to a serious extent our ability to make assured generalizations about features of the Anatolian family. These limitations will be addressed explicitly wherever appropriate.

Like other such designations (Celtic, Germanic, etc.), “Anatolian” as used here is primarily a linguistic concept. The languages named above are all assumed to be derived from a reconstructed stage “Proto-Anatolian” (PAnat.), comparable to Proto-Celtic, Proto-Germanic, and so forth, defined by certain unique shared innovations. While most of the languages of the family are in fact attested in Anatolia (Asia Minor), Luwian was also spoken in parts of Syria. On the other hand, Phrygian, although it is an Indo-European language attested from west central Anatolia, is not a member of the Anatolian subgroup as defined here.

The dearth of evidence for certain linguistic categories beyond Hittite and Luwian, and especially our poor understanding of Lydian grammar, reduces the number of assured common innovations that define Proto-Anatolian, that is, that differentiate the Anatolian subgroup from the rest of the Indo-European languages and from Proto-Indo-European. Nonetheless, one may cite a few such defining features: (1) Proto-Indo-European voiceless stops were “lenited” (probably simply voiced) between unaccented morae in Proto-Anatolian (for this formulation see Adiego Lajara 2001 and the discussion below under phonology); (2) Proto-Anatolian created a pronominal stem \*ob<sup>b</sup>ó/i-, which definitely functioned as the accented third person anaphoric pronoun (see Melchert 2009a: 156–159 with references, and below on its deictic use); (3) Proto-Anatolian developed third person subject enclitic pronouns for “unaccusative” verbs (see Garrett 1996, and below under pronouns). For a longer list of Proto-Anatolian innovations, including some more arguable features, see Melchert *forthcoming*, section 6.3.1.

I follow the majority viewpoint that the ancient Indo-European languages of Anatolia are intrusive to Asia Minor, resulting from in-migrations from somewhere to the north, via either the Balkans or the Caucasus (see further Mallory 2009 and Melchert 2011a with



references to dissenting views). Since neither the route nor the timing of the supposed entry into Anatolia can be determined with any precision, it remains an open question whether the prehistoric stage labeled Proto-Anatolian was spoken in Anatolia. Nevertheless, in the absence of contrary evidence, we may make the default assumption that the differentiation of the attested Anatolian languages was mostly associated with the spread of Indo-European speakers through Asia Minor.

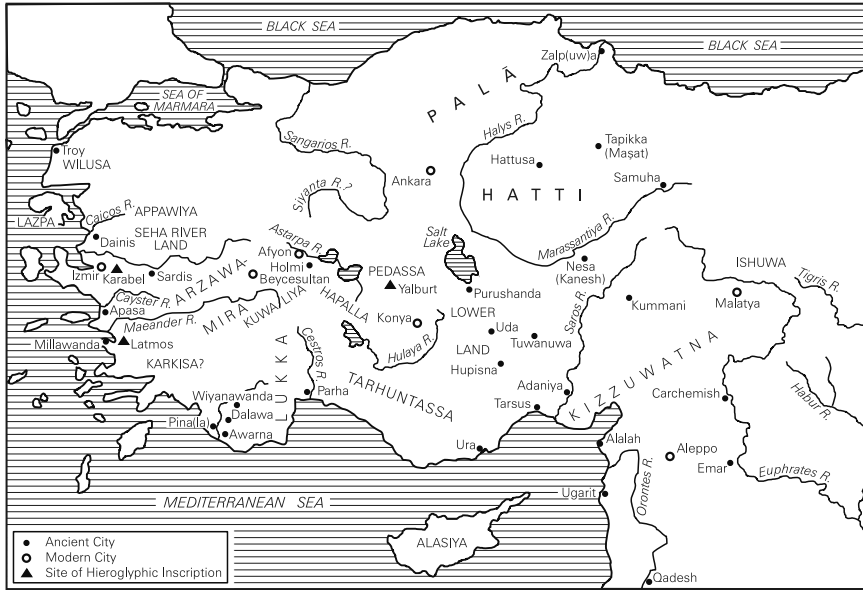
There is solid evidence for extensive prehistoric contact between Hittite and Luwian and for less intensive contact between Luwian and both Lydian and Palaic (Yakubovich 2010: Chapter Three). We may safely assume some similar contact between Lycian and Carian. Under these circumstances, attempts to describe the relationships of the Anatolian languages to each other in terms of divergence by a “family tree” model (e.g., Oettinger 1978) are less than satisfactory. Use of dialect geography presents a more realistic picture (see Melchert 2003 for detailed justification of the description given here). Luwian and Lycian undeniably share a number of innovations that justify a “Luwo-Lycian” or “Luwic” dialect group, and it is likely, but not strictly provable, that Carian belongs to this group as well. A subset of these innovations is also shared by Lydian (most notably so-called *i*-mutation, described below under nominal morphology), so that one may also speak of “western” dialects of Anatolian. On the other hand, Luwian also shares several phonological innovations with Hittite and Palaic. Since these are of a “shallower” nature than some of the features common to Luwian and the other western dialects, they might be regarded as chronologically later than the latter, but we hardly have enough facts to assert this with confidence. What does seem clear is that Luwian plays a central role in terms of shared isoglosses, as one might expect from its geographic position.

## The individual languages

### *Hittite*

The language we call Hittite was the chief administrative language of the kingdom and subsequent empire centered on the capital Hattusha in central Anatolia (see Map 3.1), which flourished from the sixteenth to the end of the thirteenth century BCE. The Hittites themselves called the language *nešili*, after the city of Nesha (also Kanish), to the southeast on the Marassantiya River (= classical Halys, modern Kızılırmak). The original locus of Hittite speakers was thus probably an area along the upper course of the Halys. They subsequently became the ruling class in Hattusha and in the Hittite kingdom, but the extent to which Hittite functioned as a spoken language beyond this group is impossible to determine. After 1350 BCE Hittite was likely in serious competition with Luwian as the spoken language even among this elite (Yakubovich 2010: Chapter Five).

Documentation of Hittite is extensive and includes a wide variety of genres: annals, treaties, letters, the state cult, therapeutic rituals, myths, and more. Our understanding of their content and of the grammar is good, but certain aspects of the phonology are obscured by limitations of the cuneiform writing system. As adopted by the Hittites, this mixed logographic-syllabic system (compare Japanese) has phonetic signs for V, CV, VC, and CVC sequences. Word-initial or word-final strings of two or more consonants and word-internal strings of three or more can only be spelled with sequences containing “empty” vowels: e.g., *pár-ah-ta* = /parxta/ ‘chased’. The system also does not fully differentiate *e*- and *i*-spellings. The resulting ambiguities leave much room for debate about the phonological interpretation of the orthography. These remarks naturally also apply to Palaic and to Luwian written in cuneiform.



MAP 3.1 ANATOLIA IN THE LATE BRONZE AGE

### *Palaic*

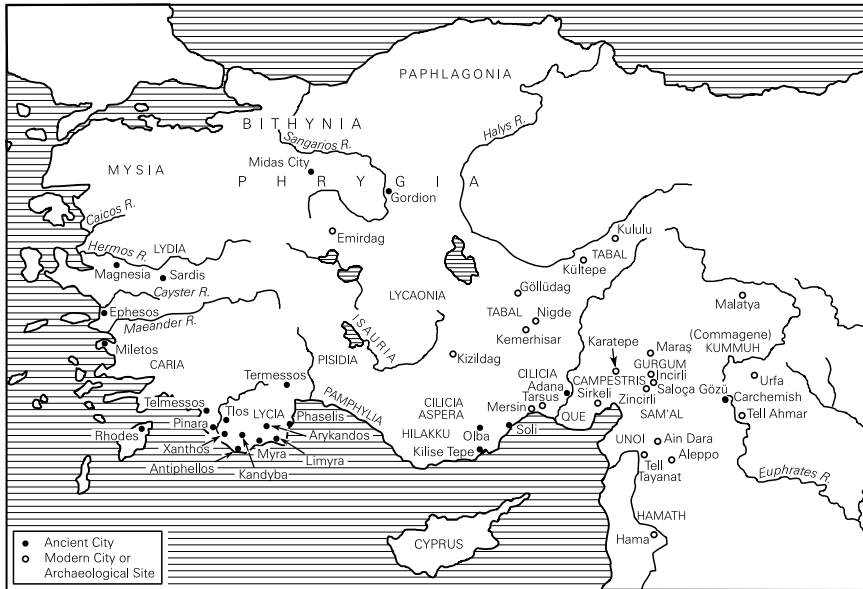
Palaic was the language of the land of Palā, almost certainly located to the northwest of the lower Halys river (Map 3.1). It is attested in a few fragmentary cuneiform texts from Hattusha (sixteenth–thirteenth centuries BCE), all dealing with the Palaic state cult, which was heavily influenced by the religion of the Hattians, a non-Indo-European people who dominated north and central Anatolia before the Hittites. Palaic is arguably the most conservative Anatolian dialect, but our evidence for it is severely limited. For further on Palaic see Carruba 1970.

### *Luwian*

Luwian is attested in cuneiform from Hattusha (sixteenth–thirteenth centuries) and in a hieroglyphic script native to Anatolia. A few hieroglyphic inscriptions date from the last Hittite kings (thirteenth century), but most hieroglyphic texts (monumental inscriptions and a few letters and economic documents on lead strips) date from the Iron Age (eleventh–eighth centuries) after the fall of the Hittite Empire, the products of various small city states and kingdoms in southern Anatolia and Syria. Traditionally, one has referred simply to Cuneiform and Hieroglyphic Luwian according to the writing system used (as followed here for convenience), but Yakubovich (2010: Chapter 1) has shown that one must distinguish between two dialects: (1) Kizzuwatna Luwian, the language of the incantation portions of Luwian rituals originating in southwestern Anatolia (Map 3.1), but as redacted by cuneiform scribes in Hittite context in Hattusha (with a few exceptions, effectively those texts collected in Starke 1985); and (2) Empire Luwian, a koine developed in and promulgated from Hattusha starting in the thirteenth century, continued in the

later Iron Age hieroglyphic texts. The “Luwianisms” scattered through Hittite cuneiform texts of the fourteenth and thirteenth centuries also belong to this dialect.

The Iron Age Luwian texts provide our richest documentation for an Anatolian language outside of Hittite, but the logographic-syllabic hieroglyphic script is even more problematic for phonological interpretation than is cuneiform. The syllabary has only V and CV signs, so that even word-final consonants must be written with an “empty” vowel, and many of the CV signs do not distinguish /a/ and /i/ vocalism. These and other practices produce spellings such as *wali+ra/i-ya-ya* for what is *probably* /warriyai/ ‘helps’. Luwian written in cuneiform helps resolve some of the resulting ambiguities, but by no means all. The Iron Age hieroglyphic texts are masterfully edited by Hawkins (2000).



MAP 3.2 ANATOLIA IN THE IRON AGE

### Lycian

Lycian is attested in close to two hundred inscriptions, nearly all on stone, from various sites in classical Lycia in southwestern Asia Minor (Map 3.2), dating from the fifth and fourth centuries BCE. These texts, written in a form of the Greek alphabet, are nearly all funerary in nature, with the exception of a few decrees and the “stele of Xanthos”, a long but poorly understood text describing the military exploits and building activities of one dynastic family. Our understanding of Lycian grammar was significantly improved by the discovery in 1973 of a decree establishing a religious cult written in Lycian, Greek, and Aramaic versions: see on the Lycian text of the “Létôon Trilingual” Laroche 1979.

A portion of the Xanthos stele text and one other inscription are written in a different dialect known conventionally as Milyan or “Lycian B” (the latter as opposed to the ordinary “Lycian A”). While the sociolinguistic relationship of the two forms of Lycian remains unclear, there is no doubt that they are dialects of the same language, despite

some claims to the contrary. Lycian A does show some unique phonological changes (\*s > h except next to consonants; \*k<sup>w</sup> > t before front vowels), but it shares important features with Lycian B that separate both from Luwian: the change of \*o to e, massive syncope of unaccented vowels, and widespread regressive vowel assimilation, among others (for arguments that Lycian is not a dialect of attested Luwian, see also Gusmani 1960).

### *Lydian*

Lydian was the language of the classical kingdom in west central Anatolia with its capital at Sardis (Map 3.2). There are now more than a hundred extant Lydian texts, but only a couple dozen are of significant length and reasonably preserved, all dating from the fifth and fourth centuries, thus roughly contemporaneous with our Lycian corpus. In the absence up to now of a Lydian-Greek bilingual text of sufficient length, our comprehension of Lydian lags far behind that of the other Anatolian languages described above, even Lycian. While we can largely parse the grammatical structure of most textual passages, many details still elude us, and our grasp of the lexicon is especially poor. As a result, we must be duly cautious in claiming that certain features are pan-Anatolian, since we do not know their status in Lydian. For all aspects of Lydian see Gusmani 1964 and subsequent supplements, especially 1986.

### *Carian, Pisidian, and Sidetic*

Carian was the language of classical Caria, in southwestern Anatolia between Lycia and Lydia (Map 3.2). A few extant texts, tentatively dated to the fourth and third centuries BCE, have been found in Caria itself, but most Carian texts, consisting of tomb inscriptions and graffiti of Carian mercenaries, are from Egypt, dating from the seventh to fifth centuries. Thanks to the discovery of a Carian-Greek bilingual text in 1986 (see Frei & Marek 1987), we can now be sure that Carian does belong to the Anatolian family of Indo-European, and it appears to share some features of the “Luwo-Lycian” dialect group. However, our understanding of Carian grammar and vocabulary remains limited, and the language cannot be regarded as fully deciphered – in particular we cannot yet analyze Carian verbal morphology, or even identify verbal forms, with confidence. Carian will thus play a very small role in the description of Anatolian grammar that follows. See on all aspects of the language Adiego 2007.

Some ten inscriptions have been discovered in Sidetic, the language of Side, a major city on the coast of Pamphylia (Map 3.2), dated very approximately to the third to second centuries BCE. Those consisting of more than personal names are largely unintelligible, and we can at present affirm only that the language is Anatolian Indo-European, belonging to the “Luwo-Lycian” dialect group. See Nollé 2001: 625–646 for the texts, but for revised readings of some letters and a tentative grammatical sketch also Pérez Orozco 2007. For Pisidian (Map 3.2) we have a mere handful of tomb inscriptions consisting of only names and patronymics, dated tentatively to the first and second centuries CE. These again can tell us only that the language is a “Luwo-Lycian” dialect of Anatolian. The very sparse evidence for both Sidetic and Pisidian leaves open the possibility that they represent late dialects of Luwian in the narrow sense. See Brixhe 1988 for the then published texts in Pisidian. We will have nothing more to say about Sidetic and Pisidian in what follows.

## PHONOLOGY

### Vowels

There was very little change in the vowels from Proto-Indo-European to Proto-Anatolian. Evidence that \*o merged with \*e in Lycian instead of with \*a as in the rest of Anatolian shows that Proto-Anatolian still maintained the PIE system of five contrasting vowel phonemes, both short and long (Rasmussen 1992 and Melchert 1992a, upheld by Hajnal 1995: 90–97): e.g., PIE/PANat. \*pedóm > Lyc. *pddē* ‘place’, PIE gen. pl. \*-óm > PANat. \*-om > Lyc. -ē. As reflected in the second example, unaccented long vowels inherited from Proto-Indo-European were shortened in Proto-Anatolian (Eichner 1973: 79 and 86<sup>xx</sup>). A PIE tautosyllabic sequence \*eh<sub>1</sub> became a Proto-Anatolian long vowel that merges with inherited long \*ē in Hittite but with *a* in Luwian and Lycian (represented here conventionally as PANat. \*ā̃): e.g., \*d<sup>h</sup>éh<sub>1</sub>ti > \*dā̃ti > Hitt. *tēzzi* ‘says’ but Lyc. *tadi* ‘puts’. Anatolian thus uniquely distinguishes between PIE \*ē and the result of tautosyllabic \*eh<sub>1</sub>.

Since Hittite maintains the PIE short diphthongs \*ay and \*oy as /ay/ and PIE \*aw and \*ow as /aw/ before coronal consonants (see Kimball 1994), these must have still been preserved at least in this environment in Proto-Anatolian. In those cases where PIE \*aw and \*ow became monophthongs in Hittite (and undoubtedly also Palaic and Luwian) they appear as a new vowel /o(:)/ distinct from inherited \*ō, which has merged with \*ā as /a:/.<sup>1</sup> As shown by Kloekhorst (2008: 35–60), vindicating the earlier claims of several scholars, the new /o(:)/ phoneme is spelled with the sign *u* in Hittite, versus /u(:)/ spelled with the sign *ú*. Prehistoric \*u is also regularly lowered to the new /o/ vowel next to \*h<sub>2</sub> or \*h<sub>3</sub>. PANat. long \*āy and \*ōy are reflected in Hitt. /a:y/, and PANat. long \*āw and \*ōw in Hitt. /a:w/ (spelled *Ca-a-ú!*). The precise timing, conditioning, and outcome of the monophthongization of PIE \*ey and \*ew and \*ēy and \*ēw remain under debate (but contra Melchert 1994a: 56 and *passim* there is no basis for supposing that the reflex of \*ey in Proto-Anatolian or Hittite was distinct from that of \*ē).

### Consonants

Evidence from Luwian and Lycian shows that Proto-Anatolian preserved the PIE three-way contrast of dorsal stops, at least for the voiceless series \*k, \*k, \*k<sup>w</sup>. However, the claim in Melchert 1989: 23–31 for an *unconditioned* contrast in Luwo-Lycian must be abandoned. A reassessment of all available evidence confirms that Anatolian is a “centum” subgroup that keeps labiovelars distinct while eventually merging the “palatal” (more accurately front velar) and velar series as velars (see for full argumentation Melchert 2012a). Before the latter merger, however, there was a conditioned palatalization in Luwo-Lycian of just the front velar series in fronting environments (\*ē, \*i, \*y and \*w): contrast, e.g., \*keyori > CLuw. *ziyar(i)* and Lyc. *sijēni* ‘lies’ with \*kēšāye/o- > CLuw. *kīšā(i)-* ‘to comb’, but \*koto- > HLuw. /kata-/ ‘enmity, opposition’ and CLuw. *kattawatnalla/i-* ‘opponent, enemy’ (cognate with Gr. κότος ‘spite’ and Skr. *sátru-* ‘enemy’). The facts for reflexes of the voiced dorsal stops are compatible with this analysis, but strict proof is prevented by the widespread *loss* of medial voiced dorsal stops in Luwo-Lycian.

There are two major changes in the stop system from Proto-Indo-European to Proto-Anatolian. First, in the absence of any compelling examples of differing reflexes, we may and should assume that the PIE voiced aspirated stops merged with voiced stops. Second, as already mentioned above, PIE voiceless stops are “lenited” (or rather simply voiced) between unaccented morae: e.g., PIE \*d<sup>h</sup>éh<sub>1</sub>ti > PANat. \*dā̃ti (i.e., \*dā̃æti) > \*dā̃ædi > Lyc. *tadi* ‘puts’, PANat. abl.-instr. \*-óti > \*-ódi > Luw. /-adi/ and Lyc. *-edi*. For this unified formulation

as a single rule see Adiego Lajara 2001, combining the two rules posited by Eichner (1973: 79–83 and 100<sup>xx</sup>) and Morpurgo Davies (1982/83). Evidence for this rule is more plentiful in Luwo-Lycian than in Hittite, where most of its effects were undone by analogy.

Before leaving the stops, we should note that Hittite scribes did not employ the cuneiform stop+vowel signs such as *ta* and *da* to indicate the phonemic contrast between voiceless and voiced stops. In intervocalic position they spelled inherited voiceless stops as geminates, while inherited voiced stops and those produced by the change just cited were spelled as single stops: thus both *pé-e-ta-an* and *pé-e-da-an* = /pe:dan/ ‘place’, contrasting with (*i-*)*ya-at-ta(-ri)* and *i-ya-ad-da(-ri)* = /yata(ri)/ ‘goes’. This practice has aroused much discussion regarding the synchronic distinctive features of the Hittite stops (see Melchert 1994a: 13–21 and compare the very different view of Kloekhorst 2008: 21–25), but for the purposes of Anatolian as a whole we may continue to speak of a voiceless vs. voiced contrast.

There were no major changes in the PIE sibilant \*s or the sonorants \*m, \*n, \*r, \*l, \*w, and \*y in Proto-Anatolian. The syllabic liquids and nasals eventually appear in the attested languages mostly as *aR* but occasionally as *uR* instead (e.g., HLuw. /t<sup>urnid</sup>/ ‘horn’ < \*k<sup>rn</sup>gid- cognate with Hitt. *karkidant-* ‘horned’), but they were likely still syllabic sonorants in Proto-Anatolian. As part of a widespread areal feature, word-initial *r-* is lacking in the attested Anatolian languages, except for extremely rare secondary instances in Luwian, Lycian, and Lydian.

As for the PIE “laryngeals”, \*h<sub>2</sub> appears in the cuneiform languages as *h* (geminate *hh* intervocalically, thus voiceless or fortis), probably a velar fricative /x/: e.g., Hitt. *hantz(a)* ‘in front’ < PIE \*h<sub>2</sub>enti (= Lat. *ante* ‘in front’ etc.) and CLuw. *hantili-* ‘foremost, first’, Hitt. *newahh-* ‘make new’ < PIE \*neweh<sub>2</sub>-. This result was subject to the Proto-Anatolian voicing rule described above, whence CLuw. 1 sg. pret. *aħa* and Lyc. *agã* ‘I made’ < PANat. \*yãya < \*yéh<sub>1</sub>-h<sub>2</sub>e versus regular CLuw. *-ħħa* and Lyc. *-xa* < PANat. \**-xa*. As shown by Kloekhorst (2006: 97–101), a sequence of \*h<sub>2</sub>w developed into a PANat. unitary labialized fricative /x<sup>w</sup>/: e.g., Hitt. /tarx<sup>w</sup>-/ ‘conquer’ spelled *tar-ħu-* and *tar-uh-* (parallel to *e-ku-* and *e-uk-* for /eg<sup>w</sup>-/ ‘drink’). In Lycian the labialized fricative turned into a labiovelar stop, represented by *q* (contra Melchert 1994a: 305–307), as shown by the name of the Storm-god, *trqqñt-* < \*trh<sub>2</sub>wñt-. PIE \*h<sub>2</sub> was lost in word-final position and before \*y and underwent assimilation in certain environments (cf. Melchert 1994a: 68–71).

PIE \*h<sub>3</sub> is reflected in word-initial position by *h-* or zero in Hittite and Luwian, under conditions that remain much debated: compare Hitt. *ħāppar* ‘transaction’ < \*h<sub>3</sub>épr̥ or \*h<sub>3</sub>ópr̥ (cognate with Lat. *opus* ‘work’ etc.) but Hitt. *arta* ‘stands (up)’ < \*h<sub>3</sub>(e)rtor. Contrary to previous opinion, Kloekhorst (2008: 838 and 946) has demonstrated that medial \*h<sub>3</sub> is preserved in the immediate environment of sonorants: CLuw. *tatarh-* ‘break’ < \*terh<sub>3</sub>- (Gr. τρώω ‘wound’) and Hitt. *walh-* ‘strike’ < \*welh<sub>3</sub>- (Gr. ἐάλω ‘perished’). For the suggestion that a sequence \*h<sub>3</sub>w is preserved in Hittite as *-hw-* (/y<sup>w</sup>/) parallel to \*h<sub>2</sub>w > *-ħhw-* = /x<sup>w</sup>/ in *lahw-* ‘pour’ (cognate with Gr. λόεω ‘bathe’ etc.) see Melchert 2011b.

Aside from secondary effects such as compensatory lengthening or assimilation, there is no compelling evidence for any direct reflex of PIE \*h<sub>1</sub> in Anatolian, *pace* Kloekhorst 2004 and 2006: 77–81 (see Melchert 2010: 152–153 and Weeden 2011: 61–68).

## Accent

The results of the PANat. “lenition” or voicing of stops and \*h<sub>2</sub> cited above suggest that the position of the inherited PIE accent was mostly retained in Proto-Anatolian. For some exceptions see Melchert 1994a: 89. For some instances of mobile accent in Hittite noun paradigms see Melchert 2012c: 180. Eichner (1986) established the synchronic accent of Lydian.

## NOMINAL MORPHOLOGY

### Nominal inflection

The attested Anatolian languages show only two numbers, singular and plural, the latter distinguishing in Old Hittite a count plural versus a collective (Melchert 2000 after Eichner 1985). Possible traces of a dual are sparse and limited to the nominative-accusative. As is well known, Anatolian also has only two attested grammatical genders, animate (usually labeled common) and neuter. It is a matter of great controversy whether Anatolian inherited the feminine gender from Proto-Indo-European and lost it (as, e.g., Armenian did) or whether its absence reflects an archaism not shared by the rest of Indo-European. One thing that is now clear is that, contrary to previous claims (Oettinger 1987, Starke 1990: 85–89, Melchert 1994b), the phenomenon of “*i*-mutation” in the western Anatolian languages is *not* a reflex of the PIE feminine gender. As first identified and described by Starke (1990: 54–85), in Luwian, Lycian, and to a lesser extent Lydian many (though by no means all!) nominal stem classes show an inserted *-i-* between the inherent stem and the endings of *just* the common gender nominative and accusative. In the case of stems in *-a-* (which mostly but not exclusively continue PIE *o*-stems) the *-i-* replaces the stem-final *-a-*. The paradigm of the Cuneiform Luwian synchronic *l*-stem *ādduwal-* ‘evil’ is nom. sg. comm. *ādduwalīš*, acc. sg. comm. *ādduwalin*, nom.-acc. sg. n. *ādduwal(=za)*, nom. pl. comm. *ādduwalinzi*, acc. pl. comm. *ādduwalinz(a)*, nom.-acc. pl. n. *ādduwala*, dat. pl. *ādduwalanz(a)*, abl.-instr. *ādduwalati*. As argued by Rieken (2005b), this pattern arose due to particular formal changes in the paradigms of some ablauting *i*-stems and was then analogically extended. This feature tells us nothing about the status of the feminine grammatical gender in Anatolian.

Old Hittite has at least eight cases: nominative, vocative, accusative, genitive, dative-locative, allative, instrumental, and ablative. The vocative is replaced by the nominative, the allative by the dative-locative, and the instrumental by the ablative by the time of New Hittite. The common gender nominative and accusative plural also merge, mostly in favor of the original accusative form. The allative (distinct from the dative-locative only in the singular) expresses goal with verbs expressing translocation (see Hoffner & Melchert 2008: 263 for examples). Its existence already in Proto-Anatolian is made likely by its use to form infinitives in multiple languages. It is indeterminate whether it is an inheritance from Proto-Indo-European or an innovation of Proto-Anatolian. Likewise, we cannot at present determine whether the merger of the dative and locative singular that took place in Hittite had already happened in Proto-Anatolian. It is also a matter of debate whether the “ergative” was a fully grammaticalized case in Proto-Anatolian (see most recently Goedegebuure 2012).

TABLE 3.1 ANATOLIAN NOMINAL ENDINGS

	Proto-Anatolian	Hittite	Palaic	Cuneiform Luwian	Hieroglyphic Luwian	Lycian	Lydian	Carian
nom. sg. comm.	*-s	-š	-š	-š	-s	zero	-š	zero
acc. sg. comm.	*-m	-n	-n	-n	-n	-V̄	-v	-n
voc. sg.	zero	zero	?	zero	zero	?	?	?
nom.-acc. sg. n.	zero	zero	zero	zero	zero	zero	-d	?
?erg. sg.	*-nts	*-nts	?	-ntiš	-ntis	-V̄ti	?	?
gen. sg.	*-os	-aš	-aš	—	-as	-V̄he	—	-š
dat.-loc. sg.	?	-i	-i/-a	-i/-a	-i/-a	-i/-eʷ	-λ	-sʷ
all. sg.	*-a	-a	-a	-a	-a	-a	?	?

	Proto-Anatolian	Hittite	Palaic	Cuneiform Luwian	Hieroglyphic Luwian	Lycian	Lydian	Carian
nom. pl. comm.	*-es	<b>-eš</b>	-eš	<b>-nzi</b>	<b>-nzi</b>	<b>-Vī</b>	?	<b>-šʷ</b>
acc. pl. comm.	*-ms	-uš	?	<b>-nz(a)</b>	<b>-nzi</b>	-Vs	?	<b>-šʷ</b>
nom.-acc. pl. n.	*-h <sub>2</sub>	zero/ <b>-a</b>	<b>-a</b>	<b>-a</b>	<b>-a</b>	<b>-a</b>	<b>-a</b>	?
?erg. pl.	*-ntes	<b>-nteš</b>	?	<b>-ntinzi</b>	<b>-ntinzi</b>	?	?	?
gen. pl.	*-om	-an	?	—	—	-ē	-av	?
dat.-loc. pl.	*-os	-aš	-aš	<b>-anz(a)</b>	<b>-anz(a)</b>	-e	<b>-av</b>	?
abl.-instr.	*-ti/-di	-z(a)	?	<b>-ati</b>	<b>-adi/-ari</b>	<b>-Vdi</b>	?	?
nom. sg. comm.	*-os	-aš	-aš	-aš	-as	-e	-aš	?
acc. sg. comm.	*-om	-an	-an	-an	-an	-ē	-ēv/-av	-n
nom.-acc. sg. n.	*-om	-an	-an	-an	-an	-ē	<b>-ad</b>	<b>-nʷ</b>
gen. sg.	*-os/-oso/-osyo	-aš	-aš	—	-as/-asi	-ehe	—	-š
dat.-loc. sg.	?	-i	-i/aʷ	-i/aʷ	-i/aʷ	-i/eʷ	-i	?
all. sg.	*-o	-a	?	?	?	-e	?	?
nom. pl. comm.	*-os	<b>-eš</b>	-aš	<b>-anzi</b>	<b>-anzi</b>	<b>-ēi</b>	?	?
acc. pl. comm.	*-oms	-uš	?	<b>-anz(a)</b>	<b>-anz(a)</b>	-es	?	?
nom.-acc. pl. n.	*-a	-a	-a	-a	-a	-a	-a	?
gen. pl.	*-om	-an	?	—	—	-ē	[-av]	?
dat.-loc. pl.	*-os	-aš	[-aš]	-aš	-as	-e	<b>-av</b>	?
abl.-instr.	*-óti/-ódi	-az(zi)	?	<b>-ati</b>	<b>-adi/-ari</b>	<b>-edi</b>	?	?

Table 3.1 summarizes the nominal case endings of Anatolian, athematic and thematic (since Anatolian has very few oxytone thematic stems, the thematic endings are given in Proto-Anatolian form, with unaccented long \*ō already shortened). Endings in boldface represent innovations, — indicates that the category has been lost, [] marks an assured ending accidentally not attested, while a question mark means that we have reason to believe the form existed but do not yet have any evidence for it.

The direct cases call for only minor comment. In addition to the expected animate nominative singular *-s*, there are a few reflexes of sonorant stems reflecting \*-ē/ōR: OHitt. *keššar*=šiš ‘his hand’ remarkably continues PIE \*ǵʰéysōr, but in most cases the PIE form has been renewed with a secondary *-s* (e.g., Hitt. and Pal. *ḫāraš* ‘eagle’ < \*ḫ<sub>3</sub>érō+s, Hitt. *ḫāsterz(a)* ‘star’ < \*ḫ<sub>2</sub>stēr+s). Lydian has generalized the pronominal neuter nominative-accusative singular ending \*(o)d to all nominal stem classes. Only Palaic preserves the animate nominative plural endings *-es* and *-ōs*, while Hittite has generalized a prehistoric *-ēs* < *-eyes*, the *i*-stem ending, and at least Luwian and Lycian, and probably Carian, have renewed the animate nominative plural ending based on the accusative plural (for details see Melchert 2009b). Old Hittite still shows some traces of the athematic nominative-accusative neuter plural ending \*-h<sub>2</sub> (e.g., *āššū* ‘goods’ = /a:sso:/ < \*-uh<sub>2</sub>), but elsewhere it has been renewed by the thematic ending *-a* < \*eh<sub>2</sub>. Hittite and Luwian also preserve some examples of lengthened-grade nominative-accusative neuter plurals such as Hitt. *widār* ‘waters’ < \*wedōr.

The history of the genitive case in Anatolian is complex. For further details of the picture presented here and references to differing views see Melchert 2012b. Hittite, Palaic, and Hieroglyphic Luwian all preserve the PIE athematic ending in the form *-os* and also a thematic ending *-o-s*. The PIE thematic genitive singular ending *-osyo* is continued by at least HLuw. /-as(s)i/ and by Carian *-š* in personal names, the latter generalized to all stem classes. The PIE thematic genitive singular *-e/oso* appears in Lyc. *-Vhe*, likewise spread to all stem classes but restricted to personal names, and probably in the Carian synchronic dative singular ending *-s*. Whether Cuneiform Luwian attests either *-e/oso* or *-osyo* is uncertain. Old Hittite, Lycian, and Lydian show reflexes of the PIE genitive



plural \*-ōm, but in Lydian the ending mostly functions as the dative-locative plural. The western Anatolian languages largely replace the genitive case with inflected forms of possessive adjectives, mostly reflecting a secondary inflection of the ending \*-osyo, but Lydian employs rather an adjective in *-la/i-* (< PIE \*-lo- with *i*-mutation), and Hieroglyphic Luwian also uses adjectives in *-iya-* < PIE \*-iyo- in this function.

As per Hajnal (1995: 98), Lycian infinitives in *-Vna* reflect the expected athematic allative ending \*-eh<sub>2</sub> (Hajnal) or less likely \*-h<sub>2</sub>e (Melchert 1994a: 324), while *-Vne* continues the analogically spread thematic ending \*-o-h<sub>2</sub> (as in Lat. *quō* ‘whither?’). It is unclear to what extent Proto-Anatolian distinguished the dative and locative singular and, if so, athematic and thematic inflection. The dative-locative plural ending \*-os attested in Hittite, Palaic, and Lycian represents an archaic form of the dative plural not yet reinforced as \*-b<sup>h</sup>(y)os or \*-mos. For more on this ending and the history of the number-indifferent ablative and instrumental in Anatolian see Melchert and Oettinger 2009.

### Nominal derivation

This topic can be treated here only in the most summary fashion. For a still largely valid overview of the presence and productivity of various PIE suffixes in Anatolian see Oettinger 1986. More focused treatments of particular stem classes are found in Weitenberg 1984, Starke 1990, and Rieken 1999a. In brief, Anatolian shows the usual mixture of archaisms and innovations and its own patterns of productivity. Root nouns are present but recessive (Rieken 1999a: 18–82). PIE \*-e/ont- is attested throughout Anatolian in a variety of functions (Melchert 2000: 58–70) but notably not in that of an active participle (see further below under non-finite verbal forms). Neuter *s*-stems of the type \*kléw-os, \*kléw-es- ‘fame’ (Skr. *śrávas-* etc.) are barely attested, if at all, but there are examples of other types (Rieken 1999a: 185–237).

Stems in *-i-* (e.g., Hitt. *arkiš* ‘testicle’ < PIE \*h<sub>2</sub>ór<sup>h</sup>is, *d(a)lugaštiš* ‘length’) are productive in Hittite but mostly recast in Luwo-Lycian in terms of *i*-mutation (see again Rieken 2005b). Stems in *-u-* (e.g., Hitt. *gēnu* ‘knee’ < \*gēnu, CLuw. *maddu* ‘wine’ < \*méd<sup>h</sup>u) are likewise reasonably productive in Hittite (Weitenberg 1984) but extremely recessive elsewhere. The well-known PIE suffix \*-ti- is modestly productive (e.g., Hitt. *išpanduzzi* ‘libation vessel’, CLuw. *arut(i)-* ‘wing’), while \*-tu- is virtually unattested. Most synchronic *l*-stems (e.g., Hitt. *išhiul* ‘binding; obligation; treaty’) are secondary from original thematic stems in \*-lo- (thus with Rieken 2008). There are relatively few reflexes of neuter *men*-stems in Hittite, but the class is very productive in Luwian, e.g., *tatariyaman-* ‘curse’ (Starke 1990: 243–299) and likely was in Lycian as well. Perhaps most notorious is the abundance in Anatolian of neuter heteroclitic stems in *-r/-n-*: see Oettinger 1986: 11–15, Rieken 1999a: 290–417, and Starke 1990: 435–525.

Primary barytone action nouns with suffix \*-o- are moderately productive, but oxytone agent nouns in \*-ó- are virtually unattested (Oettinger 1986: 19). Secondary derivatives in \*-o- are plentiful. There is a modest number of denominative stems in \*-no- and \*-to-, but exceedingly few verbal adjectives in \*-to- and none in \*-no- (cf. Oettinger 1986: 22–23). The thematic suffixes \*-iyo-, \*-lo-, and \*-ro- all enjoy some degree of productivity in Anatolian.

## PRONOUNS

### Personal pronouns

The Anatolian accented personal pronouns all continue PIE preforms, and Old Hittite maintains the PIE suppletive pattern in the first person singular and plural and second

person singular. For the second person plural we find only forms based on the PIE non-subject stem \*us- (see Katz 1998: 138–141 for competing accounts of their derivation). In New Hittite (and all the other languages for which we have evidence) the first persons and the second person plural generalize the dative-accusative, but the second person singular maintains the contrast of subject and non-subject forms (Table 3.2).<sup>2</sup>

There is considerable mutual influence between the first and second person pronouns. One defining innovation of Anatolian is the spread of the *u*-vocalism from the second person singular to the first person singular. However, the spread of the final dorsal stop from the subject form of the first person singular to the non-subject form and to both forms of the second person singular is a specifically Hittite innovation, since it is not seen in Palaic (which does not lose final stops). Likewise, reshaping of the second person plural \*us- to \*unz- based on the first person plural \*anz- < \*ŋs- is thus far attested only for Luwian. The attested enclitic non-subject forms of the first persons and the second person singular are also clearly based on PIE preforms, while the second person plural is based on the accented form, but there are signs that the Proto-Anatolian situation was more complicated, with a distinction of enclitic dative and accusative forms and subsequent differing developments of enclitic reflexives (see Yakubovich 2010: Chapter 3).

TABLE 3.2 ANATOLIAN PERSONAL PRONOUNS

	First person singular					First person plural		
	Hittite	Palaic	Hieroglyphic Luwian	Lycian	Lydian	Hittite	Cuneiform Luwian	Hieroglyphic Luwian
nom.	<i>ūk</i>	?	<i>amu</i>	<i>ēmu/amu</i>	<i>amu</i>	<i>wēš</i>	?	<i>anzanz(a)</i>
dat.-acc.	<i>ammuk</i>	?	<i>amu</i>	<i>emu</i>	<i>amu</i>	<i>anzāš</i>	<i>ānz(aš)</i>	<i>anzanz(a)</i>
enclitic	<i>-mu</i>	<i>-mu</i>	<i>-mu</i>	?	?	<i>-nnaš</i>	?	<i>-anz(a)</i>
gen.	<i>ammēl</i>	?	?	?	?	<i>anzēl</i>	?	?
abl.	<i>ammēdaz</i>	?	?	?	?	<i>anzēdaz</i>	?	?

	Second person singular			Second person plural		
	Hittite	Palaic	Hieroglyphic Luwian	Hittite	Cuneiform Luwian	Hieroglyphic Luwian
nom.	<i>zīk</i>	<i>tī</i>	<i>ti</i>	<i>šumēš</i>	?	<i>unzanz(a)</i>
dat.-acc.	<i>tuk</i>	<i>tū</i>	<i>tu</i>	<i>šumāš</i>	<i>u(n)z(aš)</i>	<i>unzanz(a)</i>
dat.-acc.	<i>-tta</i>	?	<i>-tu</i>	<i>-šmaš</i>	?	<i>-mmanz(a)</i>
enclitic						
gen.	<i>tuēl</i>	?	?	<i>šumenzan/šumēl</i>	?	?
abl.	<i>tuēdaz</i>	?	<i>tuwari</i>	<i>šumēdaz</i>	?	<i>unzari</i>

### Demonstrative and anaphoric pronouns

As noted above, one defining innovation of Proto-Anatolian is the creation of a demonstrative stem \*ob<sup>h</sup>ó/i-. This is attested as the accented third person anaphoric pronoun in Hittite, Palaic, both forms of Luwian, Lycian, and Lydian and surely had this function in Proto-Anatolian. The inflection is that of an *o*-stem (Table 3.3) in all the languages except Lydian, where it is consistently an *i*-stem. This pattern certainly has nothing to do with *i*-mutation; whether it gives a clue to the origin of the pronoun is unclear (cf. Melchert 2009a: 159 with references). Not only the Hittite nom. pl. comm. *-e* but also

the *-i-* of CLuw. *zinsi* and *zinz(a)* reflects PIE nom. pl. animate *\*-oy* (Melchert 2009b). For the Hitt. nom.-acc. pl. n. *-e* as also continuing PIE *\*-oy* see Jasanoff 2008: 144–149. The Hittite genitive singular ending *-ēl* reflects a suffix *\*-élo-* (thus with Rieken 2008: 250–251 against all others). See Goedegebuure 2007 for *apin* as the Hieroglyphic Luwian synchronic ablative-instrumental ending < PIE *\*-im* (vindicating the claim of Dunkel (1997) for such an ending) and Goedegebuure 2010 for the contrast between Luw. dat.-loc. sg. *apati* and original abl.-instr. *apadi* (attested as an adverb ‘thus’).

The enclitic anaphoric pronouns of Anatolian reflect a mixture of reflexes of PIE *\*e/o-* and *\*se/o-*, subject to various innovations, again marked in boldface. On these developments and the creation of enclitic reflexive pronouns (and the spread of reflexive *\*-ti* from Luwian to Hittite, Lycian, and Lydian) see Yakubovich 2010: Chapter 3.

TABLE 3.3 ANATOLIAN ANAPHORIC PRONOUNS

Accented third person anaphoric pronouns							
	Hittite	Cuneiform Luwian		Hieroglyphic Luwian			
nom. sg. comm.	<i>apāš</i>	<i>apāš</i>		<i>apas</i>			
acc. sg. comm.	<i>apūn</i>	<i>apān</i>		<i>apan</i>			
nom.-acc. sg. n.	<i>apāt</i>	[ <i>apa</i> ]		<i>apa</i>			
gen. sg.	<i>apēl</i>	—		<i>apas</i>			
dat.-loc. sg.	<i>apēd(an)i</i>	<i>apātti</i>		<i>apati</i>			
nom. pl. comm.	<i>apē</i>	[ <i>apīnzi</i> ]		<i>apanzi</i>			
acc. pl. comm.	<i>apūš</i>	<i>apinz(a)</i>		[ <i>apanzi</i> ]			
nom.-acc. pl. n.	<i>apē</i>	[ <i>apa</i> ]		<i>apaya</i>			
gen. pl.	<i>apenzan</i>	—		—			
dat.-loc. pl.	<i>apēdaš</i>	[ <i>apattan(z)a</i> ]		<i>apatanz(a)</i>			
abl.	<i>apēz</i>	<i>apati/apatīn</i>		<i>apin</i>			
instr.	<i>apēdand(a)/apēt</i>						

Enclitic third person anaphoric pronouns							
	Proto-Anatolian	Hittite	Palaic	Cuneiform Luwian	Hieroglyphic Luwian	Lycian	Lydian
nom. sg. comm.	<i>*-os</i>	<i>-aš</i>	<i>-aš</i>	<i>-aš</i>	<i>-aš</i>	—	<i>-aš</i>
acc. sg. comm.	<i>*-om</i>	<i>-an</i>	<i>-an</i>	<i>-an</i>	<i>-an</i>	<i>-ē(ne)</i>	<i>-av</i>
nom.-acc. sg. n.	<i>*-od</i>	<i>-at</i>	<i>-at</i>	<i>-ata</i>	<i>-ada</i>	<i>-ede</i>	<i>-ad</i>
dat. sg.	<i>*-soy</i>	<i>-šše/-šši</i>	<i>-tu</i>	<i>-tu</i>	<i>-tu</i>	<i>-i</i>	<i>-mλ</i>
nom. pl. comm.	<i>*-oy</i>	<i>-el/-at</i>	<i>-aš</i>	<i>-ata</i>	<i>-ada</i>	?	?
acc. pl. comm.	<i>*-oms</i>	<i>-uš/-aš</i>	?	<i>-aš</i>	<i>-ada</i>	?	?
nom.-acc. pl. n.	<i>*-oy</i>	<i>-el/-at</i>	<i>-e</i>	<i>-ata</i>	<i>-ada</i>	?	?
dat. pl.	<i>*-sm-os</i>	<i>-šmaš</i>	?	<i>-mmaš</i>	<i>-mmanz(a)</i>	<i>-ñne</i>	<i>-mš</i>

Determining the Proto-Anatolian system of deictic pronouns is made difficult by our lack of evidence for the languages other than Hittite: we have almost no data for Cuneiform Luwian, or for anything but near deixis in Lycian, Lydian, and Carian. Furthermore, some cognate stems have very different deictic values in the various languages. The overview here reflects Melchert 2009a, but much remains indeterminate. The only thing that seems clear is that the Proto-Anatolian marker of near deixis was *\*k̄o-* (compare Lat. *citrā* ‘on this side’ etc. < *\*ki-*), reflected in Hitt. and Pal. *kā-*, CLuw. and HLuw. *zā-*, and

probably Carian *s(a)-*. The inflection is entirely that of an *a*-stem, except for the aberrant and as yet unexplained Hitt. nom.-acc. sg. n. *kī*. As shown by Goedegebuure (2002/2003), Hittite has a three-way deictic contrast like that of Lat. *hic, iste, and ille*: *kā-* for near the speaker, *apā-* for near the addressee, and *aši/uni* (< PIE \*é/ó- plus particle \*-i) for far deixis. In attested Hieroglyphic Luwian, *apa-* covers the range of both *iste* and *ille*, but there is indirect evidence that at least Cuneiform Luwian once also had far-deictic \*é/ó- (Melchert 2009a: 152). Very confusingly, however, Lyd. *es-* < \*é/ó- marks *near* deixis, and Lyc. *ebe-* does as well. It is quite unclear just which stem(s) marked far deixis in Proto-Anatolian and whether it had a system with a two- or three-way contrast.

### Enclitic possessive pronouns

Emphatic pronominal possession is expressed throughout the history of Hittite by the genitive of the personal pronouns and the accented anaphoric pronoun, preceding their head noun. In Old Hittite, unemphatic pronominal possession is marked by enclitic pronouns (more properly pronominal adjectives) attached to the head noun and agreeing with it in gender, number, and case: e.g., nom. sg. comm. *attaš=miš* ‘my father’, nom.-acc. sg. n. *šaḫeššar=šummet* ‘our fortress’, dat.-loc. pl. *ḫaššaš=šaš* ‘for his offspring’ (see Hoffner & Melchert 2008: 137–141). However, there were no specific forms for the nominative-accusative plural neuter or the ablative, and the nominative-accusative singular neuter and the instrumental forms were used respectively by suppletion: *šākuwa=šmet* ‘their eyes’ and *iššaz=(š)mit* ‘from their mouths’. By New Hittite, use of the enclitic possessives was no longer part of the synchronic grammar, as shown by the fact that copyists misunderstood the forms in *-et* and *-it* and misused them for cases like the vocative and dative-locative singular. Unemphatic possession in New Hittite was conveyed by use of the appropriate enclitic dative pronoun: *nu=šši KUR=SÚ GUL-un* (conj. + ‘him’ (dat.) ‘land’ (obj.) ‘I struck’) = ‘I struck his land’. It is likely that Cuneiform Luwian attests a few examples of enclitic possessives (Carruba 1986), and the category is probably to be reconstructed for Proto-Anatolian (thus Garrett 1991–1993: 160–161).

TABLE 3.4 ANATOLIAN INTERROGATIVE-RELATIVE PRONOUN

	Hittite	Cuneiform Luwian	Hieroglyphic Luwian
nom. sg. comm.	<i>kuiš</i>	<i>kuīš</i>	REL- <i>is</i>
acc. sg. comm.	<i>kuin</i>	<i>kuīn</i>	REL- <i>in</i>
nom.-acc. sg. n.	<i>kuit</i>	<i>kuī</i>	REL- <b>anz(a)</b>
gen. sg.	<i>kuēl</i>	—	—
dat.-loc. sg.	<i>kuedani</i>	<i>kuwatti</i>	REL- <i>ati</i>
nom. pl. comm.	<i>kuēš/kuīēš</i>	<i>kuīnzi</i>	REL- <i>inzi</i>
acc. pl. comm.	<i>kuiuš</i>	[ <i>kuīnz(a)</i> ]	REL- <b>inzi</b>
nom.-acc. pl. n.	<b><i>kuē</i></b>	?	REL- <b>ya</b>
gen. pl.	[ <i>kuenzan</i> ]	—	—
dat.-loc. pl.	<i>kuedaš</i>	?	?
abl.	<i>kuēz</i>	<i>kuwatti(n)</i>	REL- <i>adi</i>

### Interrogative, relative, and indefinite pronouns

Anatolian uses the PIE stem \*k<sup>w</sup>o/i- as both the interrogative and relative pronoun – there is no trace of PIE relative \*Hyo-. The interrogative-relative was clearly inflected as an *i*-stem in the nominative and accusative, subject to very minor innovations: besides the

Hittite and Luwian forms in Table 3.4 note also Pal. *kui-*, Lyc. *ti-*, Lyd. *qi-*, and Carian *xi-*. In the other cases the interrogative shows the same endings as the anaphoric demonstrative stems in Hittite, and our very limited evidence suggests the same for the other languages. Hittite attests the occasional use of the bare interrogative stem *kui-* as an indefinite, but the latter is usually expressed by a form with an added particle: nom. sg. comm. *kuiški*, acc. sg. comm. *kuinki*, gen. sg. *kuēlka*, etc., and the other languages display similar creations: CLuw. *kuišḫa*, Lyc. *tike* and *tise*.

## VERBAL MORPHOLOGY

### Verbal categories

The Anatolian verb system is markedly different in certain basic features from that of the oldest stages of Indo-Iranian and Greek and thus from that traditionally reconstructed for PIE. As with the absence of the feminine gender, this glaring discrepancy has been the object of long and heated controversy: is it to be explained by loss of categories in Anatolian, by shared innovations in the rest of the Indo-European languages, or by a combination of the two? There is not remotely a consensus on this important issue, and what follows necessarily represents a particular viewpoint. One should not, however, overstate the matter: many aspects of the Anatolian verb are reassuringly familiar and not subject to serious dispute. One should also distinguish between the presence/absence of categories and the presence/absence or relative productivity of particular formal means of expressing a given category.

The Anatolian verb is inflected for the expected three persons and for two numbers, singular and plural. There is a contrast between active and medio-passive voice, and the latter attests the functions seen in other older Indo-European languages. Oppositional middles indicate “subject affectedness”: (1) reciprocal (Hitt. med.-pass. *zahhanta* ‘they strike each other’ vs. act. *zahhanzi* ‘they strike’); (2) self-interest (Hitt. med.-pass. *ušneškatta* ‘offers for sale’ < \*‘pledges in return for payment’ vs. act. *ušneškemi* ‘I pledge’; and (3) passive (Hitt. med.-pass. *ḫullattati* ‘was struck’ vs. act. *ḫullet* ‘struck’). The passive use is rare in Old Hittite but becomes highly productive by New Hittite. Delimiting the functions of *media tantum* is notoriously more difficult, but Hittite shows most of the well-known types: (1) body care/grooming (Hitt. *wēšta* ‘wears’), (2) change of body posture (Hitt. *kitta(ri)*, Pal. *kītar*, CLuw. *ziyar(i)*, Lyc. *sijēni/sitēni* ‘lies (down)’), (3) non-translational motion (Hitt. *wēḫatta(ri)* ‘spins (intr.)’), (4) translational motion (Hitt. *iyatta(ri)* ‘walks’), (5) mental states (Hitt. *lēlaniyatta* ‘becomes angry’), and (6) “spontaneous” events/changes of state (Hitt. *kīša(ri)* ‘happens; becomes’).

The Anatolian verb displays only two moods: indicative and imperative. There are no obvious reflexes of the PIE optative or subjunctive. In Hittite the function of the optative to express the non-factual (what is possible, contrary to fact, or desired) is filled by the particle *mān/man*, a secondary use of the subordinating conjunction *mān* ‘when(ever), if’, itself derived from a rare PIE interrogative-relative stem \**mo-* (compare Hitt. *maši-* ‘how/as many’, Toch. B *māksu* ‘who, which’). For a likely indirect trace of the PIE subjunctive in Hittite see Jasanoff 2012, but compare Oettinger 2007.

The Anatolian verb has two tenses, present (which also serves for the future) and preterite. Hittite does famously develop a periphrastic perfect with *ḫark-* ‘hold, have’ and *ēš-* ‘be’ and the past participle, comparable to similar formations in modern western European languages; see Boley 1984 and, on the correlation of auxiliary selection and unaccusativity, Garrett 1996: 102–106. As described in summary fashion below, Anatolian attests

virtually all of the suffixes used to form the “present”, that is, imperfective, stem in “Core Indo-European”. It is quite uncertain, however, whether Anatolian inherited a true aspectual contrast of imperfective (“present”) vs. perfective (“aorist”). The evidence presented in Melchert 1997 for reflexes of such a contrast is suggestive, but far from compelling, and one may also suppose that Anatolian inherited the various present-forming suffixes merely as markers of Aktionsart and that the development of a true aspectual system is a common innovation of the non-Anatolian languages; for a well-reasoned presentation of this viewpoint see Strunk 1994.

Synchronically, most, if not all verbs are “monothematic”, showing a single stem that may be understood as perfective or imperfective according to context. However, as shown by evidence that is plentiful in Hittite and limited but clear in the other languages, Anatolian developed the means to optionally mark a finite verb as explicitly imperfective, most often by addition of a reflex of the PIE suffix *\*-ské/o-*, but also for some verbs of *\*-enh<sub>2</sub>i-* or *\*-s(e/o)-* (whose PIE antecedents have not been fully elucidated). This “marked” imperfectivity may be realized as various Aktionsarten, depending on the context and the lexical semantics of the base verb: progressive, iterative, durative, habitual, inceptive, or distributive (see Hoffner & Melchert 2008: 318–323 for illustrations from Hittite).

There is no Anatolian category corresponding to the PIE perfect, generally agreed to have expressed an attained state, whether one defines it as an aspect (e.g., Meier-Brügger 2000: 155) or an Aktionsart (e.g., DiGiovine 1996: 273). Once again, explanations for this absence differ radically, as discussed further in the next section.

### Verbal inflection: active

The Hittite finite verb follows in the active voice one of two inflectional patterns, labeled after the respective present first-person singular endings the “*mi*-conjugation” and the “*hi*-conjugation” (Table 3.5). Evidence from Luwian (see Morpurgo Davies 1979) suggests that this system is Proto-Anatolian, although examples from Lycian and Palaic are extremely sparse and not beyond question. Even in Hittite there was much mutual influence between the two conjugations, and in the other attested languages the difference may have been reduced to just the present third singular indicative.

TABLE 3.5 HITTITE VERBAL INFLECTION

Active indicative				
<i>mi</i> -conjugation			<i>hi</i> -conjugation	
Present				
	Singular	Plural	Singular	Plural
1	<i>-mi</i>	<i>-weni/-wani</i>	<i>-h<sub>2</sub>he/-h<sub>2</sub>hi</i>	<i>-weni/-wani</i>
2	<i>-ši</i>	<i>-teni/-ttani</i>	<i>-tti/-š<sub>2</sub>ti</i>	<i>-(š)teni/-(š)tani</i>
3	<i>-zzi</i>	<i>-anzi</i>	<i>-e/-i/-āi</i>	<i>-anzi</i>
Preterite				
	Singular	Plural	Singular	Plural
1	<i>-un/-anun</i>	<i>-wen</i>	<i>-h<sub>2</sub>hun</i>	<i>-wen</i>
2	<i>-š/-tta</i>	<i>-ten</i>	<i>-tta/-š<sub>2</sub>ta</i>	<i>-(š)ten</i>
3	<i>-t(a)</i>	<i>-ēri/-ar</i>	<i>-š/-š<sub>2</sub>ta</i>	<i>-ēri</i>

Active imperative				
<i>mi</i> -conjugation			<i>hi</i> -conjugation	
	Singular	Plural	Singular	Plural
1	<i>-allu</i>	<i>-weni</i>	<i>-allu</i>	<i>-weni</i>
2	<i>zero/-t</i>	<i>-tten</i>	<i>zero/-i</i>	<i>-tten</i>
3	<i>-ttu</i>	<i>-antu</i>	<i>-u</i>	<i>-antu</i>

Medio-passive indicative				
Present			Preterite	
	Singular	Plural	Singular	Plural
1	<i>-(h)ha(ri), -(h)haḥari</i>	<i>-wašta(ti)</i>	<i>-(h)hat(i), -(h)haḥat(i)</i>	<i>-waštat(i)</i>
2	<i>-(t)ta(ri), -(t)tati</i>	<i>-t(t)uma(ri)</i>	<i>-at, -(t)tat(i)</i>	<i>-(d)dumat</i>
3	<i>-a(ri), -(t)ta(ri)</i>	<i>-anta(ri)</i>	<i>-at(i), -(t)tat(i)</i>	<i>-antat(i)</i>

Medio-passive imperative		
	Singular	Plural
1	<i>-hḥaru, -haḥaru</i>	<i>-waštati</i>
2	<i>-ḥut(i)</i>	<i>-(d)dumat(i)</i>
3	<i>-aru, -(t)taru</i>	<i>-antaru</i>

The singular endings of the *mi*-conjugation manifestly match the primary and secondary endings standardly reconstructed for the present-aorist system in PIE. While only Hittite has 1 sg. pres. *-mi* (see below on the other languages), Palaic, Luwian, and Lydian also attest to 2 sg. pres. *-si*, and these plus Lycian attest to 3 sg. pres. *-ti* (only Hittite assibilates \*t before syllabic \*i). Hitt. 1 sg. pret. *-un* continues athematic \*-m̄, while *-anun* reflects thematic \*-om renewed by the athematic ending. Lyd. 1 sg. pret. *-v* continues at least \*-om (with syncope) and perhaps also \*-m̄. Older Hittite still preserves *-š* as the 2 sg. pret. ending of vocalic stems (e.g., *iēš* ‘you did’), but consonant stems have already adopted *-ta* from the *hi*-conjugation (e.g., *ēpta* ‘you took’). Hittite and Palaic show 3 sg. pret. *-V-t* but *-C-ta*, the latter with a genuine “prop-vowel” (Hitt. 3 sg. pret. *e-ku-ut-ta* ‘drank’ = /eg<sup>w</sup>ta/, not †e-ku-ut). On the other hand, the consistent 3 sg. pret. ending *-(t)a* of Luwian and *-te/-de* of Lycian after a vowel as well as a consonant reflects renewal of the third person preterite active endings by those of the medio-passive (thus with Yoshida 1993).

The Anatolian first person plural active endings show a *-w-*, recalling the PIE first person dual endings, but also an *-n* matching the Greek dialectal 1 pl. act. *-μεν* (cf. Eichner 1975: 79). Hitt. 1 pl. pret. *-wen* and Lyd. *-wv* reflect a virtual \*-wen, to which was added the usual mark of the present \*-i to form 1 pl. pres. \*-weni continued by Hitt. and Pal. *-weni* and CLuw. *-wanni/-unni*. On the model of the first person plural in \*-wen(i), Anatolian created a matching second person plural \*-ten(i), extending PIE \*-te, attested in Hittite, Palaic, and Luwian. In posttonic position \*-weni and \*-teni regularly became

Hitt. *-wani* and *-t(t)ani* (Melchert 1994a: 137–138 after Cowgill). The ending *-t(t)ani* has no connection whatever to the *\*-th<sub>2</sub>ene* seen in Skr. *-thana*, contra Eichner 1975: 79 and others. The present third person plural ending of Hittite *-anzi* and *-anti* elsewhere probably reflect PIE athematic *\*-enti* and *\*-nti* as well as thematic *\*-onti*, but it is difficult to establish this beyond doubt. Only Pal. 3 pl. pret. *-anta* may preserve a trace of the expected secondary ending *\*-ent/-nt/-ont* (but it cannot be excluded that Pal. *-anta* reflects medio-passive *\*-onto*, as in Luwian and Lycian; see Yoshida 1993: 34, note 21). In Hittite, the third person plural preterite ending of the *hi*-conjugation has been generalized to all verbs, usually in the form *-ēr* < *\*-ēr*, but there are a few examples of *-ar* < *\*-r*. Much less certain is whether Lyd. 3 pl. pret. *-rś/-ris* continues PIE *\*-rs*, matching Skr. *-ur* (see Melchert 2004a: 147, note 14, on the synchronic ending). The Anatolian preterite shows no credible traces of the PIE “augment” *\*é-*.

The present singular endings of the *hi*-conjugation strongly resemble those of the PIE perfect and can in fact be straightforwardly derived from endings matching the latter plus the *\*-i* particle that characterizes the present indicative: PIE *\*-h<sub>2</sub>ey* > OHitt. *-h<sub>2</sub>ey* and PIE *\*-ey* > OHitt. *-e*, already mostly leveled to *-h<sub>2</sub>i* and *-i* after the majority of present endings ending in *-i*. The leveling has already occurred in Hitt. 2 sg. pres. *-tti* for *\*-tte* < *\*-th<sub>2</sub>ey*. The third singular ending is renewed partially in Hittite and totally elsewhere by an ending *-āi* spread from monosyllabic stems where it results from contractions (thus Luw. *-āi* and probably Pal. *-ai* and Lyc. *-e*). An entirely unsolved problem is the generalized 1 sg. pres. *-wi* of Luwian, which is surely matched by Lyd. *-Cu/-Vw* (with regular apocope) and probably also by Lyc. *-u*.

Endings matching the perfect also appear in 1 sg. pret. Pal. *-h<sub>2</sub>ha*, CLuw. *-h<sub>2</sub>ha*, HLuw. *-ha*, and Lyc. *-xa/-ga* < *\*-h<sub>2</sub>e* (Hitt. *-h<sub>2</sub>un* has been reshaped after *mi*-conjugation *-un*). However, Hitt. 3 sg. pret. *-š* obviously cannot be connected with the PIE perfect, and the ending *-tta* in Luwian and *-te* in Lycian is again a renewal by the medio-passive ending *\*-to* (Yoshida 1993: 31–33) that tells us nothing about the Proto-Anatolian state of affairs.

It is extremely unlikely that the Anatolian *hi*-conjugation is to be derived from the PIE perfect, or vice versa (for trenchant criticism of attempts at the former see Jasanoff 2003: 7–17), but just what the true relationship of the two categories is remains to be determined. Jasanoff (2003), elaborating earlier publications, has made a strong case for both a “*h<sub>2</sub>e*-present” and a “*\*h<sub>2</sub>e*-aorist” in PIE, but many issues remain unresolved. On the status of *\*ó/é* and *\*ó/zero* ablaut in root *\*h<sub>2</sub>e*-verbs see Melchert 2013 versus Kloekhorst 2012. Regarding the ablaut pattern of *\*h<sub>2</sub>e*-presents in *-i-* (Hitt. 3 sg. pres. *dāi*, 3 pl. pres. *tianzi* ‘put(s)’) compare the differing analyses of Kimball (1998), Jasanoff (2003: 98–107), and Kloekhorst (2008: 807 and *passim*). For the source of the 3 sg. pret. ending *-š* of the *hi*-conjugation and of the *-š-* that appears in the second person plural of *hi*-verbs in *-i-*, one may contrast the proposals of Jasanoff (2003: 174–197 and 2012: 118–119), Kloekhorst (2007), and Melchert (2015).

The status of the perfect in Anatolian also remains disputed. Jasanoff (2003: 11 and 37 and 117–118) argues that Hitt. *wewakk-* ‘demand’ and *mēm/a-* ‘speak’ continue reduplicated perfects. See also Forssman 1994: 103 for Hitt. *šipand-* ‘libate; sacrifice’ < *\*spēspōnd-* (vs. *išpant-* ‘idem’ < *\*spēnd/spōnd-*). But even if one or more of these verbs reflects a perfect in *formal* terms (which is far from assured), none demonstrably shows the “attained state” meaning assumed for “Core Indo-European”. Oettinger (2001: 80–83 and 2006: 37–42) views the perfect (along with the present type of Skr. *dadhāti* and the intensive type of Skr. *várvariti*) as a post-Anatolian development of a PIE reduplicated present that appears in mostly de-reduplicated form in the Anatolian *hi*-conjugation. Hajnal (1999: 8–25) sees rather the Anatolian *hi*-conjugation and the post-Anatolian perfect as



separate developments of a PIE “proto-perfect”, the perfect having been formally influenced by inherited intensive presents. A reasoned choice between these competing scenarios is not yet possible.

The imperative is marked in the third person by replacing the *-i* of the indicative endings with *-u*, a clearly Proto-Anatolian feature, as shown by the equation of 3 sg. imp. Hitt. *ēšdu*, Pal. *āšdu*, CLuw. *āšdu*, HLuw. /a:stu/, and Lyc. *esu* ‘let be’. In Hittite this *-u* was spread to the first person singular voluntative in *-allu*, whose prehistory remains unclear. The exhortative first person plural in *-weni* is formally identical to the indicative and can be identified only by context; whether it continues a PIE subjunctive is impossible to determine. The second person singular imperative is, as expected, usually the bare verbal stem with a zero ending. One finds the reflex of the PIE ending *\*-dʰi* in the Hitt. 2 sg. imp. *īt* ‘go!’ (cognate with Skr. *thi* and Gr. *ἴθι*) and in verbs with the suffix *-nu-*: 2 sg. imp. *warnut* ‘burn!’. The second person plural imperative continues PIE augmentless forms with secondary endings. Since the preterite indicative also lacks the “augment”, one can distinguish second person plural preterites and imperatives only by context.

### Verbal inflection: medio-passive

The endings of the PIE medio-passive are also ultimately related to those of the perfect and the Anatolian *hi*-conjugation, but this relationship is a pre-PIE matter. Anatolian incontrovertibly inherited a formally and functionally distinct medio-passive voice. As demonstrated by Yoshida (1990), the present medio-passive was marked in Proto-Anatolian at least in the third person by endings in *\*-r* (a feature shared with Italo-Celtic and Tocharian). This ending was lost by regular sound change, totally in Lycian, but only partially in Hittite, Luwian, and Palaic (the situation in Lydian is unclear). In the present of attested Hittite, the *r*-endings were restored, reinforced by the particle *-i* of the present active (Table 3.5). The details of the developments in Palaic and Luwian remain unclear due to the paucity of evidence (see Yoshida 1990: 115–117). In Lycian we find a few medio-passive forms marked by a nasal ending (1 sg. pres. *sixani* ‘I lie’, 3 sg. pres. *sijēni/sitēni* ‘lies’, 1 sg. pret. *axagā* ‘I became’, for which see Melchert 1992b). The source of this inflection is quite uncertain.

In the preterite and also in the second person singular and first person plural of the present, the Hittite medio-passive shows inherited endings enlarged by a particle *-ti* (reduced to *-t* in the preterite after Old Hittite, surely because verbal endings in *-i* were strongly associated with the present tense). As already seen by Neu (1968: 145), this element *-ti* ([*-di*]) must be a variant of the reflexive particle *\*-ti*, which in the latter function appears assimilated as Hitt. *-z(a)*. See for details of the derivation Yakubovich 2010: 199–205 and Yoshida 2010: 237–238. Late Hieroglyphic Luwian attests preterite medio-passives ending in *-si* (Oshiro 1993: 53–54 and Rieken 2004), which must somehow continue a reflexive use of *\*-soy* (attested as such in Palaic), but the appearance of HLuw. *-si* in this function so late raises serious issues of relative chronology (cf. Yakubovich 2010: 201–202).

As in the active, the imperative is marked in the third persons by replacing *-i* with *-u* (cf. also CLuw. 3 sg. imp. *āyaru* and HLuw. 3 sg. imp. /i<sup>h</sup>iyaru/ ‘let it become’), and at least in Hittite this was also spread to the first person singular. Again as in the active the first person plural exhortative in Hittite is formally identical to the indicative, and the second person plural is identical with the preterite indicative. The Hittite 2 sg. imp. ending *-hut(i)* is of uncertain origin, though it is likely that the *-t(i)* is again the reflexive particle [*-di*].

### Verbal stem formation

As indicated above, whether or not Anatolian inherited a fully developed aspectual contrast of imperfective versus perfective, it certainly does preserve evidence for virtually every type of PIE “present”-forming stem. The following summary agrees on most substantial points with the presentation in Oettinger 1979, with the important updates in Oettinger 2002. We find the following familiar primary formations: (1) root presents (PIE \*h<sub>1</sub>ésti, \*h<sub>1</sub>s-énti > Hitt. *ēšzi*, *ašanzi*, CLuw. *āšti*, \*ašanti, Lyc. *esi*, \*ahñti ‘is, are’), (2) acrostatic “Narten” presents (PIE \*wēkti, \*wēkñti > Hitt. *wēkzi*, *wēk(k)anzi* ‘demand(s)’), (3) barytone and oxytone media tantum (PIE \*k<sub>1</sub>éyor > CLuw. *ziyar(i)*, Lyc. *sijēni* ‘lies’ and renewed \*k<sub>1</sub>éytor > Hitt. *kitta(ri)*, Pal. *kītar*, and Lyc. *sitēni*; PIE \*t<sub>1</sub>duk<sub>1</sub>or > Hitt. *tuqqāri* ‘is visible; important’), (4) nasal infix presents (e.g., PIE \*h<sub>2</sub>u-né-g-ti, \*h<sub>2</sub>u-n-g-énti > Hitt. *hūnikzi*, *hūninkanzi* ‘wound(s)’ – see Shatskov 2006 on the Hittite stem formation – and PIE \*d<sup>h</sup>w<sub>1</sub>-né-h<sub>1</sub>-ti, \*d<sup>h</sup>w<sub>1</sub>-n-h<sub>1</sub>-énti > Hitt. *d(u)warnezzi*, *d(u)warnanzi* ‘breaks’ – see Kloekhorst 2008: 907 on the Hittite stem formation), (5) verbs with the suffix \*-n(e)u- (extremely productive, as in Hitt. *hūinu-* and CLuw. *hūinu(wa)-* ‘run’ and Lyc. *qanuwe-* ‘cause to be slain’), (6) presents in \*-ye/o- with a full-grade root (PIE \*k<sub>1</sub>érp-ye-ti, \*k<sub>1</sub>érp-yonti > Hitt. *karpīēzzi*, *karpīyanzi* ‘lift(s)’), Lyc. *fa-korfīd* ‘undertakes’), (7) iterative presents in \*-éye/o- and a zero-grade root (virtual \*(s)tubh-éye/o- > CLuw. *dūpiti*, *dupainti* and Lyc. *tubīdi*, *tubeiti* ‘strike(s)’), and (8) causative presents in \*-éye/o- and an *o*-grade root (PIE \*lowkéye/o- > Hitt. *lukkezzi*, *lukkanzi* ‘kindle(s)’ – see Oettinger 2002: xx).

Simple thematic presents existed in Anatolian: HLuw. *tama-* (/damma-/) ‘build’ with 3 sg. pres. AEDIFICARE+MI-ri+i = /dammari/ < \*démh<sub>2</sub>-eti – where the rhotacized ending reflects \*-di voiced from \*-ti between unaccented morae – and Hitt. *šuwezzi*, *šuwanzi* ‘push(es), reject(s)’ < \*suh<sub>1</sub>-éti, \*suh<sub>1</sub>-ónti (Oettinger 1979: 297). They are, however, startlingly rare vis-à-vis “Core Indo-European”: for one account of this situation see Jasanoff 1998. As already discussed above, “*i*-presents” in Hittite belong to the *hi*-conjugation, with an ablaut pattern that is disputed. The status of “*u*-presents” is unclear, with Hittite attesting one clear case in the *mi*-conjugation (*tarhuzzi*/*taruḫzi* ‘conquers’ < \*térh<sub>2</sub>u-ti – see Kloekhorst 2006: 98–101 and 2008: 835–838) and one in the *hi*-conjugation (*lāhui* ‘pours’ < \*lōh<sub>3</sub>w-ey – see Melchert 2011b). There are numerous reduplicated stems in Anatolian, but they have not yet received a systematic treatment, and the possible relationships of the various types to those of PIE remain to be elucidated.

PIE denominative formations are also well attested in Anatolian. Those in \*-ye/o- are very productive in all the languages and hardly need illustration. We also find factitives in \*-eh<sub>2</sub>, which as shown by Hittite originally belonged to *hi*-conjugation (Jasanoff 2003: 139): OHitt. 3 sg. pres. *šuppiyahhi* ‘purifies’. In New Hittite there is transfer to the *mi*-conjugation (*šuppiyahzi*), and this shift is complete in the other languages (e.g., Lyc. *prñnawati* ‘builds’), but the secondary nature of this inflection is betrayed by the lack of voicing of the *-t-* of the ending (correct Hajnal 1995: 162, note 82, contra Melchert 1994a: 69). Anatolian also attests the extended factitive suffix \*-eh<sub>2</sub>-ye/o- represented by the Hittite type of *armāizzi*, *armanzi* ‘impregnate(s)’ (beside *armahh-*) and Lyc. *xttadi*, *xttaiti* ‘harm(s)’.

Continuants of PIE root aorists are assured (Eichner 1975: 82): e.g., Hitt. *kuerzi*, *kuranzi* = CLuw. *k(u)warti*, \*kuranti ‘cut’ (Skr. *ákar* ‘made’) and Hitt. *tēzzi* ‘says’ = Lyc. *tadi* ‘puts’ (Skr. *ádhāt* ‘put’). Whether there are any attested reflexes of the PIE sigmatic aorist is uncertain. Anatolian tells us nothing about the PIE status of the thematic or reduplicated aorist.

### Non-finite verbal forms

Proto-Anatolian clearly formed productive neuter verbal nouns with a suffix *\*-wer/-wen-*. Hittite evidence suggests that these originally were inflected according to the “proterokinetic” pattern, in idealized form nom.-acc. R(é)-w<sub>ṛ</sub>, gen. R(zero)-wén-s. Hittite productively generalizes the post-vocalic form *-war* of the nom.-acc. and the gen. sg. *-waš* < *\*-wen-s*. There are, however, some traces of post-consonantal *\*-w<sub>ṛ</sub>* and a generalized zero grade of the suffix: Hitt. *ḫenkur*, *ḫenkunaš* ‘offering’. Hittite no longer has a full paradigm of this verbal noun, but the old abl.-instr. *\*-wen-ti* functions as the infinitive in *-wanzi*, and the endingless locative in *\*-wen* appears as the “supine” in *-wan* (see below for its use). The other languages appear to have generalized the zero-grade form of the suffix, based on the use of the allative of the paradigm as the infinitive in Palaic, Luwian, and Lycian: Pal. *-una*, Luw. *-una*, and Lyc. *-na* < PANat. *\*-una* < *\*-uneh<sub>2</sub>* (as per Hajnal 1995: 98, the more common Lyc. *-ne* reflects analogical influence of the thematic allative ending *\*-oh<sub>2</sub>*). The other attested Hittite infinitive in *-anna* is a specific Hittite innovation, representing the allative of the Hittite verbal abstract in *\*-eh<sub>2</sub>-t<sub>ṛ</sub>*, *\*-eh<sub>2</sub>-tn-*: virtual *\*-eh<sub>2</sub>-tn-eh<sub>2</sub>*. Lydian infinitives have not yet been identified with confidence. Brief mention should also be made of Hieroglyphic Luwian gerundives in *-min(a)*, thus far attested only in predicational function in nominal sentences with the meaning ‘(X is) to be \_\_ed’ (Melchert 2004b).

One of the great surprises of Anatolian from the viewpoint of other older Indo-European languages is that, as intimated earlier, there is hardly any trace of verbal adjectives in *\*-to-* and none in *\*-no-*. Instead, one finds in the function of a past participle (that is, one expressing an attained state) derivatives in *\*-e/ont-*. This is the productive form in Hittite, predictably usually with a passive sense for transitive verbs (e.g., *kunant-* ‘killed’), but merely resultative for intransitives (e.g., *uwant-* ‘come, arrived’). Also unsurprising is that in generic use (i.e., without an expressed object) the participles of transitive verbs can have an active sense (e.g., *adant-* ‘having eaten’, *paḫ(ḫa)šnuwant-* ‘protecting, on guard’). How one is to explain this use versus the active *processual* meaning of *\*-e/ont-* in “Core Indo-European” has yet to be decided.

In Luwian and Lycian there are mere relics of the past participle in *\*-e/ont-* (e.g., CLuw. *walant-/ulant-* and Lyc. *lāta-* ‘dead’). The productive past participle is of the form *-mma/i-* = *-me/i-* (with the ubiquitous *i*-mutation). Despite its enduring popularity, the derivation of this suffix from a PIE medio-passive participle cannot possibly be correct. One must insist on the fact that the function of this participle is *precisely* the same as that of Hitt. *-ant-*: it does *not* express an ongoing activity “being Xed”, but rather an attained state, predictably usually passive for transitive verbs: e.g., CLuw. *dūpaimma/i-* ‘struck’, *awimma/i-* ‘come, arrived’. The usual exceptions apply to transitive verbs, and we find HLuw. */adamma/i-/* ‘having eaten’ and */uwamma/i-/* ‘having drunk’. Given this meaning and the rampant productivity of neuter verbal nouns in *\*-men* precisely in Luwo-Lycian, there should be no doubt that these past participles originate in secondary thematic derivatives with possessive sense: virtual *\*-o-mn-o-* (Melchert 2014: 206–207).

### ADVERBS AND ADPOSITIONS

There is no truly productive means of forming manner adverbs in Anatolian. Both Hittite and Luwian show use of the nominative-accusative plural neuter of adjectives in this function: e.g., Hitt. *karši* (and renewed *karšaya*) ‘frankly’ < *karši-* ‘bare, unadulterated; frank’ and CLuw. *wāšu* ‘well’. Hittite does show a limited productive use of adverbs in *-ili* (e.g., *duddumili* ‘silently’ < *duddumili-* ‘silent, mute’), almost certainly the nominative-accusative plural neuter of adjectives in *\*-ili-* (compare Lat. *-ilis*).

What calls for extended comment is the Anatolian system of “local adverbs”, attested as free-standing adverbs, adpositions, and preverbs. Starke (1977: 127–180) established that in Old Hittite there were several pairs of such adverbs, one expressing location and the other direction. However, there are signs that the Old Hittite situation is in certain respects innovative vis-à-vis Proto-Anatolian. E.g., the form of the Hittite *hi*-verb *āppāi*, *āppianzi* ‘step back; be finished’ argues that the Proto-Anatolian form of the adverb ‘back’ was \**ópi*, as attested in CLuw. *āppi*/HLuw. *api*, and that Hitt. *āppa* ‘back; again’ has been reshaped after other directional adverbs that came to end in *-a* in Hittite (such as *anda* ‘into’ < \**endo*).

Despite repeated claims to the contrary, there is no evidence that most Anatolian local adverbs represent frozen case forms of original substantives. As described in Melchert 2009c, the use of enclitic possessives with local adverbs in Old Hittite is in most cases merely analogical to new postpositions that manifestly *are* case forms of nouns, such as *tapušz(a)* ‘beside’ < *tapuš-* ‘side’. The one notable exception is the family of Hitt. *šēr* ‘above; over’ and *šarā* ‘up’ (with plentiful cognates elsewhere in Anatolian). Here we truly are dealing with an archaic noun referring to a height: compare Gr. *ῥίον* ‘mountain peak’. Much work remains to be done in establishing the Proto-Anatolian antecedents of the respective systems of the attested languages.

## SYNTAX

### Morphosyntax

The use of the nominal cases in Anatolian is mostly unremarkable in Indo-European terms. We find the “dative of disadvantage” with persons: e.g., OHitt. <sup>L1</sup>SAGI (subj.) LUGAL-*i* (dat. sg.) NINDA.GUR<sub>4</sub>.RA (obj.) *ēpzi* ‘takes’ = ‘The drink-server takes the leavened bread from the king’. Due to the formal merger of the dative and locative, the dative-locative could analogically also be used to express ‘place from which’ with inanimate objects, in competition with the ablative: OHitt. *kardī=šmi=at* ‘your heart’ (dat. sg.) + ‘it’ (nom.-acc. sg.) *dāḥḥun* ‘I have taken’ = ‘I have taken it (the sickness) from your heart’. Hieroglyphic Luwian shows the same use of the dative-locative beside the ablative-instrumental (see KARKAMIŠ A6, §§27–28, Hawkins 2000: 125). Due to a reanalysis of a dative of goal beside an infinitive expressing purpose with motion verbs, Anatolian came to use the dative-locative for objects of infinitives besides the accusative: see Melchert 2012 (2014). On the controversial issue of “split ergativity” see Goedegebuure 2012 with references to previous analyses.

To express ‘begin to \_\_’, Hittite uses either a combination of the reflexive particle *-z(a)*, the finite verb *ēp-* ‘take’, and the infinitive, or the finite verb *dai-* ‘put’ (secondarily also *tīya-* ‘step’) plus the so-called supine in *-wan* (see Hoffner & Melchert 2008: §§25.20 and 25.37–38). Hittite also famously develops a “serial” or “phraseological” construction that combines a finite form of *pai-* ‘go’ or *uwa-* ‘come’ with another finite verb in a single clause. The precise meaning, synchronic structure, and historical source of this construction have elicited a variety of analyses. For a recent discussion see Koller 2013 with ample references to previous treatments.

### Configurational syntax: word order

The unmarked, functionally neutral word order of all the second-millennium languages is clearly SOV. Lydian also is unmistakably SOV. Lycian has manifestly innovated, but the paucity of diagnostic examples has made it hard to determine just what the “basic”

Lycian word order is; for a tentative claim of VSO see Garrett 1994: 30–32. One must emphasize the notion of “unmarked” or “basic” word order, because other orders of major constituents are by no means rare, and recent intensified study of Hittite order has revealed even more variation than previously acknowledged. In trying to make sense of this variety, one must distinguish two separate though interrelated issues: (1) how do various configurations correlate with discourse structure (topic, focus, anaphora); and (2) which orders are licensed by the syntax, and which by the phonology? Significant progress has been made in addressing these questions, but much work remains to be done. One complicating factor is that it is now clear that for Hittite one cannot use as *primary* data material from mythological texts or ritual incantations, since word order in these may show influence from Hattic or Hurrian (see Rizza 2007 and Rieken 2011). A systematic overview is impossible here and would be almost immediately outdated in any case. In addition to the summary in Hoffner and Melchert 2008: 406–409 (flawed by the uncritical use of material from translation literature), one may get a sense of what has been achieved and how much remains to be done by consulting Bauer 2011, Goedegebuure 2013 (with references to previous analyses), and Sideltsev *forthcoming*.

One famous feature of Anatolian syntax is that most clauses in connected discourse are linked to a preceding clause by “sentence-initial” conjunctions to which are attached sequences of enclitics (unemphatic anaphoric pronouns plus a variety of “particles”) by what is known as “Wackernagel’s Law”. For persuasive arguments that said conjunctions are actually interclausal see Agbayani and Golston 2010. The beginnings of this system are certainly Proto-Anatolian, but several facts argue that it was there only embryonic and that each of the Anatolian languages has elaborated it in its own fashion. First, there are differences in the order of the clitics: in Hittite the singular dative clitics follow third person nominative or accusative clitics (*n=at=mu/ši* conj. + nom.-acc. sg. n. ‘it’ + ‘to me’/‘to him’), but in Luwian the order is the reverse (HLuw. /a/=wa/=m(u)/=an/ conj. + quotative marker + ‘to me’ + acc. sg. comm. ‘it’). Second, different morphemes are used for the same function: Hittite uses *=ma* to mark contrastive focus, while Palaic and Luwian employ *=ppa* in this function. Third, the same morpheme appears in different functions: PIE \**nú* ‘now’ still appears as such in Pal. *nū*, but it has begun to be grammaticalized as a clause-linking conjunction (Carruba 1970: 65), a development completed in Hittite, while the Lydian cognate *-in* (with apocope of the original vowel and then anaptyxis) comes at the end of the clitic sequence and likely has an asseverative value (that it is cognate cannot be doubted, contra Gusmani 1964: 133).

Nevertheless, some features can tentatively be attributed to Proto-Anatolian. The only clause-linking conjunction that may be Proto-Anatolian is \**oh*<sub>1</sub> (cf. Dunkel 2007: 57), the frozen instrumental of the PIE \**é/ó-* anaphoric pronoun, attested as *a-* in Hittite (rare), Palaic, and both forms of Luwian, and in extended Lyd. *ak*, in all cases with a prosecutive sense, roughly ‘and then’. Contrary to the claim of Watkins (1963: 13–16) and others, the Old Hittite set of clause-linking conjunctions *šu*, *ta*, and *nu* cannot be compared with OIr. *se*, *to*, and *no*. As shown by Weitenberg (1992: 327), OHitt. *šu* is used only with the preterite, while *ta* is used with the present-future. This distribution supports the etymology of *šu* < PIE \**h*<sub>1</sub>*su* (Zimmer apud Dunkel 2007: 56–57) and of *ta* < instr. \**toh*<sub>1</sub> (Rieken 1999b: 86). In any case, all three of the cited conjunctions are creations of Hittite. Of the various “local particles” that follow the anaphoric pronouns in the clitic sequence, that most likely to have existed already in Proto-Anatolian is \**-te*, which probably indicated movement from one spatial domain across a boundary into another (Josephson 1995:

171), attested in Pal. and Luw. *-tta*, Lyd. *-(i)t-*, and Hitt. *-(a)šta*, the last via metanalysis (Josephson 1972: 419).

### Configurational syntax: subordinate clauses

Anatolian has conditional and temporal subordinating clauses, using a variety of conjunctions built on interrogative-relative stems (see Hoffner & Melchert 2008: 414–423 for an overview of these in Hittite). Differences in detail among the various languages leave in doubt the status of such clauses in Proto-Anatolian. Anatolian lacks “final clauses” and uses parataxis (Hitt. *nu*, Lyc. *me*): e.g., Hitt. *takku LÚ-an pahḫueni kuiški peššiezzi n=aš aki* ‘if anyone makes a man fall into a fire, so that he dies’.

Our understanding of Hittite and Anatolian relative clauses has changed dramatically in the last decade, and ongoing research promises to bring further changes in the next few years. Contrary to a long-standing consensus (Garrett 1994, following Hale 1987), Goedegebuure (2009) and Huggard (2011) have shown that Hittite does not have overt “wh-movement”; the surface order of wh-interrogative and relative clauses is governed by considerations of focus. Furthermore, there is also counterevidence to the claim that Hittite distinguishes “determinate” from “indeterminate” preposed relative clauses in that the former must be preceded by at least one full constituent, while the latter occur only clause-initially (not counting clause-linking conjunctions and any attached clitics); see Held 1957, Hale 1987: 46–49, and Garrett 1994: 45–49. Note the two indeterminate relatives preceded by focused constituents in the following New Hittite text (KBo 5.4 Ro 33–34):

*namma ANA<sup>a</sup>UTU-ŠI kuiš LÚ.KÚR [n=aš=tta] LÚ.KÚR ēšdu tuk=ma kuiš LÚ.KÚR ANA*  
 further to His Majesty who enemy conj.=he=you enemy let be you=focus who enemy to  
<sup>a</sup>UTU-ŠI=ya=aš LÚ.KÚR  
 His Majesty=also=he enemy  
 ‘Furthermore, whoever is an enemy to His Majesty shall be an enemy to you, while-  
 whoever is an enemy to you is an enemy to His Majesty.’

While preposed correlative relative clauses of the type just shown are statistically by far the most common type in Hittite, Probert (2006) showed that Old Hittite also has a type of embedded relative clause. Current research reveals that Hittite also has relative clauses embedded *between* constituents of the main clause, as well as “free relatives”, while postposed restrictive and non-restrictive relatives are far more common than previously acknowledged.

### LEXICON

Tischler (1979) demonstrated that contrary to a widespread perception (see, e.g., Kronasser 1956: 219), at least two-thirds of Hittite’s core vocabulary that is known is either inherited from PIE or built on inherited material. The misperception that a majority of the Hittite lexicon is non-Indo-European is due to the nature of our textual evidence, much of which deals with the state cult and rituals. Since Hittite religion and ritual were heavily influenced by Hattic and Hurrian, unsurprisingly much of the vocabulary for the relevant personnel and apparatus is borrowed. Nor is it unusual that a significant portion of the words for Anatolian flora and fauna are loanwords. This has nothing to do with core vocabulary, which is securely Indo-European. Our knowledge of the core vocabulary of the other Anatolian languages is woefully inadequate, but what is known suggests the same for them.

## THE POSITION OF ANATOLIAN

The notable differences between the Anatolian languages and the grammar reconstructed for PIE on the basis of the other subgroups has elicited a variety of responses. Some scholars have attributed them nearly entirely to changes in Anatolian (mostly losses, hence the overly simplistic label “Schwundhypothese”): see Pedersen 1938 and Eichner 1975. Others have argued that Anatolian is in Stammbaum terms a “sister” of reconstructed PIE, one that preserves many archaic features of a much earlier proto-language labeled “Indo-Hittite”, from which PIE had markedly diverged: thus Sturtevant 1933 and Lehrman 1998. In the 1960s a number of scholars proposed that the reconstruction of PIE itself, especially the verbal system, needed to be drastically revised, on the basis of data not only from Anatolian but also from other subgroups: see the very diverse views of Adrados 1963, Meid 1963, Neu 1968, and Watkins 1969.

After more than two decades during which a compromise between these diametrically opposed views seemed unlikely, there has since developed a growing consensus that the non-Anatolian languages did participate in *some* common innovations setting them apart from Anatolian, but that these innovations are neither particularly numerous nor for the most part profound. Nor do they preclude that Anatolian shared in certain post-PIE innovations with particular subgroups (cf. Puhvel 1994). Debate now centers on just how many revisions must be made to PIE to account for the features of Anatolian – or, to phrase it differently, just which features of “Core Indo-European” (= “Restindogermanisch”) actually represent common innovations not shared by Anatolian. For three individual viewpoints one may compare Rieken 2009, Oettinger 2013–14, and Melchert 2013 *forthcoming*, but one should not conclude that these by any means represent the full spectrum of opinion.

## FURTHER READING

In addition to the works already cited, one may consult the following: Popko (2008) offers a useful overall summary; see for an overview of Hittite Rieken 2005a and for grammatical sketches of all the languages the relevant chapters in Woodard 2004. Hoffner and Melchert (2008) describe Hittite synchronic grammar *in extenso*. Tischler (2001) furnishes a reliable basic dictionary for Hittite. The two major dictionaries for Hittite (Friedrich-Kammenhuber-Hoffmann [1975–] and the *CHD* [1980–]) will take decades to complete. Tischler (1977–2016), Puhvel (1984–), and Kloekhorst (2008) provide three different viewpoints on Hittite etymology. For Hittite texts, editions, bibliography, onomastics, and much more, one should consult the invaluable online “Hethitologie Portal Mainz”: <http://www.hethport.uni-wuerzburg.de/HPM/index.html>.

## NOTES

- 1 In the special case of word-final accented \*-óm(s), PIE \*o does appear as Hitt. /o:/, thus acc. sg. comm. *ku-u-un* = /ko:n/< \*kóm and comm. acc. pl. *ku-u-uš* = /ko:s/< \*kóms: cf. Kloekhorst 2008: 54–57.
- 2 For the Hieroglyphic Luwian first and second person plural as ending in /-ants/, not /-unts/, see Yakubovich 2010: 65–68.

## REFERENCES

- Adiego, Ignacio J. 2007, *The Carian Language*, Brill, Leiden–Boston  
 Adiego Lajara, Ignacio-Javier 2001, “Lenición y acento en protoanatolico,” in: Caruba, Onofrio & Meid, Wolfgang (eds.), *Anatolisch und Indogermanisch: Anatolico*

- e Indoeuropeo*, Institut für Sprachen und Literaturen der Universität Innsbruck, Innsbruck: 11–18
- Adrados, Francisco R. 1963, *Evolución y estructura del verbo indoeuropeo*, Instituto Antonio de Nebrija, Madrid
- Agbayani, Brian & Chris Golston 2010, “Second-position is first-position: Wackernagel’s Law and the role of clausal conjunction,” *Indogermanische Forschungen* 115: 1–21
- Bauer, Anna 2011, “Verberststellung im Hethitischen,” in: Krisch, Thomas & Lindner, Thomas (eds.), *Indogermanistik und Linguistik im Dialog: Akten der XIII. Fachtagung der Indogermanischen Gesellschaft vom 21. bis 27. September 2008 in Salzburg*, Reichert, Wiesbaden: 39–48
- Boley, Jacqueline 1984, *The Hittite Hark- Construction*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck
- Brixhe, Claude 1988, “La langue des inscriptions épichorique de Pisidie,” in: Arbeitman, Yoël (ed.), *A Linguistic Happening in Memory of Ben Schwarz*, Peeters, Louvain-la-Neuve: 131–155
- Carruba, Onofrio 1970, *Das Palaische. Texte, Grammatik, Lexikon*, Harrassowitz, Wiesbaden
- Carruba, Onofrio 1986, “Die 3. Per. Sg. des Possessivpronomens im Luwischen,” in: Hoffner Jr., Harry A. & Beckman, Gary M. (eds.), *Kaniššwar: A Tribute to Hans G. Güterbock*, The Oriental Institute, Chicago: 49–52
- CHD = Güterbock, Hans G., Hoffner Jr., Harry A. & van den Hout, Theo 1980, *The Hittite Dictionary of the Oriental Institute of the University of Chicago*, Oriental Institute, Chicago
- DiGiovine, Paolo 1996, *Studio sul perfetto indoeuropeo II*, Il Calamo, Rome
- Dunkel, George 1997, “B. Delbrück and the Instrumental-Ablative in \*-m,” in: Crespo, Emilio & García-Ramón, José Luis (eds.), *Berthold Delbrück y la sintaxis indoeuropea hoy. Actas del Coloquio de la Indogermanische Gesellschaft. Madrid, 21–24 de septiembre de 1994*, Ediciones de la UAM–Reichert, Madrid–Wiesbaden: 63–83
- Dunkel, George 2007, “Chips from an Aptologist’s workshop I,” in: Nussbaum, Alan J. (ed.), *Verba Docenti: Studies in Historical and Indo-European Linguistics Presented to Jay H. Jasanoff by Students, Colleagues, and Friends*, Beech Stave, Ann Arbor–New York: 53–61
- Eichner, Heiner 1973, “Die Etymologie von heth. mehur,” *Münchener Studien zur Sprachwissenschaft* 31: 53–107
- Eichner, Heiner 1975, “Die Vorgeschichte des hethitischen Verbalsystems,” in: Rix, Helmut (ed.), *Flexion und Wortbildung. Akten der V. Fachtagung der Indogermanischen Gesellschaft. Regensburg, 9.-14. September 1973*, Reichert, Wiesbaden: 71–103
- Eichner, Heiner 1985, “Das Problem des Ansatzes eines urindogermanischen Numerus ‘Kollektiv’ (‘Komprehensiv’),” in: Schlerath, Bernfried (ed.), *Grammatische Kategorien. Funktion und Geschichte. Akten der VII. Fachtagung der Indogermanischen Gesellschaft. Berlin, 20.-25. Februar 1983*, Reichert, Wiesbaden: 135–169
- Eichner, Heiner 1986, “Die Akzentuation des Lydischen,” in: Schindler, Jochem & Peters, Martin (eds.), *Festgabe für Manfred Mayrhofer (= Die Sprache 32)*, Verlag der Wiener Sprachgesellschaft, Vienna: 7–21
- Forssman, Bernhard 1994, “Zu hethitisch *šipand-* und *išpant-*,” in: Rasmussen, Jens (ed.), *In honorem Holger Pedersen. Kolloquium der Indogermanischen Gesellschaft vom 26. bis 28. März 1993 in Kopenhagen*, Reichert, Wiesbaden: 93–106
- Frei, Peter & Marek, Christian 1987, “Die karisch-griechische Bilingue von Kaunos,” *Kadmos* 36: 1–89



- Friedrich, Johannes, Kammenhuber, Annelies & Hoffmann, Inge 1975, *Hethitisches Wörterbuch*, 2nd ed., Winter, Heidelberg
- Garrett, Andrew 1991–1993, “A note on the morphosyntax of lycian and anatolian possession,” *Die Sprache* 35: 155–161
- Garrett, Andrew 1994, “Relative clause syntax in lycian and hittite,” *Die Sprache* 36: 29–69
- Garrett, Andrew 1996, “Wackernagel’s Law and unaccusativity in Hittite,” in: Halpern, Aaron & Zwicky, Arnold (eds.), *Approaching Second: Second Position Clitics and Related Phenomena*, CSLI, Stanford: 85–133
- Goedegebuure, Petra 2002/2003, “The Hittite 3rd person/distal demonstrative *aši* (*uni, eni*, etc.),” *Die Sprache* 43: 1–32
- Goedegebuure, Petra 2007, “The Hieroglyphic Luwian demonstrative ablative-instrumentals *zin* and *apin*,” in: Archi, Alfonso & Francia, Rita (eds.), *VI Congresso Internazionale di Ittologia. Roma, 5–9 settembre 2005*, Consiglio Nazionale delle Ricerche, Rome: 319–334
- Goedegebuure, Petra 2009, “Focus structure and q-word questions in Hittite,” *Linguistics* 47: 945–969
- Goedegebuure, Petra 2010, “The Luwian demonstratives of place and manner,” in: Singer, Itamar (ed.), *ipamati kistamati pari tumatimis. Luwian and Hittite Studies Presented to J. David Hawkins on the Occasion of his 70th Birthday*, Institute of Archaeology, Tel Aviv University, Tel Aviv: 76–94
- Goedegebuure, Petra 2012, “Split-ergativity in Hittite,” *Zeitschrift für Assyriologie* 102: 270–303
- Goedegebuure, Petra 2013, “Hittite noun phrases in focus,” in: Jamison, Stephanie W. & Melchert, H. Craig & Vine, Brent (eds.), *Proceedings of the 24th Annual UCLA Indo-European Conference*, Hemen, Bremen: 27–45
- Gusmani, Roberto 1960, “Concordanze e discordanze nella flessione nominale del licio e del luvio,” *Rendiconti del Istituto Lombardo di Scienze e Lettere. Classe di lettere e scienze morali e storiche* 94: 497–512
- Gusmani, Roberto 1964, *Lydisches Wörterbuch*, Winter, Heidelberg
- Gusmani, Roberto 1986, *Lydisches Wörterbuch. Ergänzungsband*. Lieferung 3, Winter, Heidelberg
- Hajnal, Ivo 1995, *Der lykischer Vokalismus*, Leykam, Graz
- Hajnal, Ivo 1999, “Vergleichende Rekonstruktion im Wandel. Vom Paradigma zur Funktion,” *vorgesehen für: Akten des Gyarmathi-Symposiums, Göttingen, 19.-21. November 1999*. [[http://sprawi.uibk.ac.at/files/hajnal/b37\\_vergleichende\\_rekonstruktion.pdf](http://sprawi.uibk.ac.at/files/hajnal/b37_vergleichende_rekonstruktion.pdf), seen 2 July, 2014]
- Hale, Mark 1987, “Notes on Wackernagel’s Law in the language of the Rigveda,” in: Watkins, Calvert (ed.), *Studies in Memory of Warren Cowgill (1929–1985)*, de Gruyter, Berlin–New York: 38–50
- Hawkins, J. David 2000, *Corpus of Hieroglyphic Luwian Inscriptions. Volume 1: Inscriptions of the Iron Age*, de Gruyter, Berlin
- Held, Warren H. Jr. 1957, *The Hittite Relative Sentence*, Linguistic Society of America, Baltimore
- Hoffner, Harry A. & H. Craig Melchert 2008, *A Grammar of the Hittite Language. Part 1: Reference Grammar*, Eisenbrauns, Winona Lake
- Huggard, Mattyas 2011, “On *Wh*-(Non)-movement and internal structures of the Hittite preposed relative clause,” in: Jamison, Stephanie W.; Melchert, H. Craig & Vine, Brent (eds.), *Proceedings of the 22nd Annual UCLA Indo-European Conference*, Hemen, Bremen: 83–104

- Jasanoff, Jay 1998, "The thematic conjugation revisited," in: Jasanoff, Jay; Melchert, H. Craig & Oliver, Lisi (eds.), *Mír Curad: Studies in Honor of Calvert Watkins*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck: 301–316
- Jasanoff, Jay 2003, *Hittite and the Indo-European Verb*, Oxford University Press, Oxford–New York
- Jasanoff, Jay 2008, " \*-bhi, \*-bhis, \*-ōis: Following the trail of the PIE instrumental plural," in: Rasmussen, Jens & Olander, Thomas (eds.), *Internal Reconstruction in Indo-European: Methods, Results, and Problems. Section Papers from the XVI International Conference on Historical Linguistics. University of Copenhagen, 11th–15th August, 2003*, Museum Tusulanum, Copenhagen: 137–149
- Jasanoff, Jay 2012, "Did Hittite have *si*-imperatives?," in: Sukač, Roman & Šefčik, Andrej (eds.), *The Sound of Indo-European 2. Papers on Indo-European Phonetics, Phonemics and Morphophonemics*, Lincom Europa, Munich: 116–132
- Josephson, Folke 1972, *The Function of the Sentence Particles in Old and Middle Hittite*, Skriv Service AB, Uppsala
- Josephson, Folke 1995, "Directionality in Hittite," *Analecta Indoeuropaea Cracoviensia* 2: 165–176
- Katz, Joshua T. 1998, *Topics in Indo-European Personal Pronouns*, Harvard University [PhD dissertation]
- Kimball, Sara 1994, "The IE short diphthongs \*oi, \*ai, \*ou and \*au in Hittite," *Die Sprache* 36: 1–28
- Kimball, Sara 1998, "On the origins of Hittite verbs of the type *tēhli, dāi*," in: Jasanoff, Jay; Melchert, H. Craig & Oliver, Lisi (eds.), *Mír Curad: Studies in Honor of Calvert Watkins*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck: 335–344
- Kloekhorst, Alwin 2004, "The preservation of \**h*<sub>1</sub> in Hieroglyphic Luwian: Two separate *a*-signs," *Historische Sprachforschung* 117: 26–49
- Kloekhorst, Alwin 2006, "Initial laryngeals in anatolian," *Historische Sprachforschung* 119: 77–108
- Kloekhorst, Alwin 2007, "The Hittite 2pl.-ending *-šten(i)*," in: Archi, Alfonso & Francia, Rita (eds.), *VI Congresso Internazionale di Ittitologia. Roma, 5–9 settembre 2005*, Consiglio Nazionale delle Ricerche, Rome: 493–500
- Kloekhorst, Alwin 2008, *Etymological Dictionary of the Hittite Inherited Lexicon*, Brill, Leiden–Boston
- Kloekhorst, Alwin 2012, "Hittite *ā/e*-ablauting Verbs," in: Melchert, H. Craig (ed.), *The Indo-European Verb. Proceedings of the Conference of the Society for Indo-European Studies. Los Angeles, 13–15 September 2010*, Reichert, Wiesbaden: 151–160
- Koller, Bernhard 2013, "Hittite *pai-* 'go' and *uwa-* 'come' as restructuring verbs," *Journal of Historical Linguistics* 3: 77–97
- Kronasser, Heinz 1956, *Vergleichende Laut- und Formenlehre des Hethitischen*, Winter, Heidelberg
- Laroche, Emmanuel 1979, "L'inscription lycienne," *Fouilles de Xanthos* 6: 49–127
- Lehrman, Alexander 1998, *Indo-Hittite Redux*, Paleograph, Moscow
- Mallory, James P. 2009, "The anatolian homeland hypothesis and the anatolian neolithic," in: Jamison, Stephanie; Melchert, H. Craig & Vine, Brent (eds.), *Proceedings of the 20th Annual UCLA Indo-European Conference*, Hempten, Bremen: 133–162
- Meid, Wolfgang 1963, *Die indogermanischen Grundlagen der altirischen absoluten und konjunkten Verbalflexion*, Harrassowitz, Wiesbaden
- Meier-Brügger, Michael 2000, *Indogermanische Sprachwissenschaft*, 7e völlig neubearbeitete Auflage, de Gruyter, Berlin–New York

- Melchert, H. Craig 1989, "New luvo-lycian isoglosses," *Historische Sprachforschung* 102: 23–45
- Melchert, H. Craig 1992a, "Relative chronology and anatolian: The vowel system," in: Beekes, Robert; Lubotsky, Alexander & Weitenberg, Jos (eds.), *Rekonstruktion und Relative Chronologie. Akten der VIII. Fachtagung der Indogermanischen Gesellschaft Leiden, 31. August – 4. September 1987*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck: 41–53
- Melchert, H. Craig 1992b, "The middle voice in lydian," *Historische Sprachforschung* 105: 189–199
- Melchert, H. Craig 1994a, *Anatolian Historical Phonology*, Rodopi, Amsterdam–Atlanta
- Melchert, H. Craig 1994b, "The feminine gender in anatolian," in: Dunkel, George E.; Gisela Meyer, Salvatore Scarlata & Christian Seidl (eds.), *Früh-, Mittel- Spätindogermanisch. Akten der IX. Fachtagung der Indogermanischen Gesellschaft vom 5. bis 9. Oktober 1992 in Zürich*, Reichert, Wiesbaden: 231–244
- Melchert, H. Craig 1997, "Traces of a PIE. Aspectual contrast in anatolian?," *Incontri Linguistici* 20: 83–92
- Melchert, H. Craig 2000, "Tocharian plurals in *-nt-* and related phenomena," *Tocharian and Indo-European Studies* 9: 53–75
- Melchert, H. Craig 2003, "The dialectal position of lycian and lydian within anatolian," in: Giorgieri, Mauro; Mirjo Salvini, Marie-Claude Trémouille & Pietro Vannicelli (eds.), *Licia e Lidia prima dell'Ellenizzazione. Atti del Convegno internazionale Roma, 11–12 ottobre 1999*, Consiglio Nazionale delle Ricerche, Rome: 265–272
- Melchert, H. Craig 2004a, "Second thoughts on *\*y* and *\*h<sub>2</sub>* in Lydian," in: Mazoyer, Michel & Casabonne, Olivier (eds.), *Studia Anatolica et Varia: Mélanges offerts au professeur René Lebrun*, vol. II, Kubaba, Paris: 139–150
- Melchert, H. Craig 2004b, "Hieroglyphic Luvian Verbs in *-min(a)*," in: Hyllested, Adam; Anders Richardt Jørgensen, Jenny Helena Larsson & Thomas Olander (eds.), *Per Aspera ad Asteriscos: Studia Indogermanica in honorem Jens Elmegård Rasmussen sexagenarii Idibus Martiis anno MMIV*, Institut für Sprachen und Literaturen der Universität Innsbruck, Innsbruck: 355–362
- Melchert, H. Craig 2009a, "Deictic pronouns in anatolian," in: Yoshida, Kazuhiko & Vine, Brent (eds.), *East and West: Papers in Indo-European Studies*, Hempen, Bremen: 151–161
- Melchert, H. Craig 2009b, "The animate nominative plural in luvian and lycian," in: Nedoma, Robert & Stifter, David (eds.), *\*h<sub>2</sub>nr: Festschrift für Heiner Eichner*, Harrassowitz, Wiesbaden: 112–117
- Melchert, H. Craig 2009c, "Local adverbs in Hittite: Synchrony and diachrony," *Language and Linguistic Compass* 3/2: 607–620
- Melchert, H. Craig 2010, "Spelling of initial /a-/ in hieroglyphic Luwian," in: Singer, Itamar (ed.), *ipamati kistamati pari tumatimis. Luwian and Hittite Studies presented to J. David Hawkins on the Occasion of his 70th Birthday*, Institute of Archaeology, Tel Aviv University, Tel Aviv: 147–158
- Melchert, H. Craig 2011a, "Indo-Europeans," in: Steadman, Sharon R. & McMahon, Gregory (eds.), *The Oxford Handbook of Ancient Anatolia. 10,000–323 B.C.E.*, Oxford University Press, Oxford: 704–716
- Melchert, H. Craig 2011b, "The PIE verb 'to pour' and medial *\*h<sub>3</sub>* in anatolian," in: Jamison, Stephanie W.; Melchert, H. Craig & Vine, Brent (eds.), *Proceedings of the 22nd Annual UCLA Indo-European Conference*, Hempen, Bremen: 127–132

- Melchert, H. Craig 2012a, "Luvo-lycian dorsal stops revisited," in: Sukač, Roman & Šefčík, Andrej (eds.), *The Sound of Indo-European 2. Papers on Indo-European Phonetics, Phonemics and Morphophonemics*, Lincom Europa, Munich: 206–218
- Melchert, H. Craig 2012b, "Genitive case and possessive adjective in anatolian," in: Orioles, Vincenzo (ed.), *Per Roberto Gusmani. Studi in ricordo* vol. II. *Linguistica storica e teorica*, Forum, Udine: 273–286
- Melchert, H. Craig 2012c, "Hittite 'Heteroclite' s-Stems," in: Cooper, Adam I.; Rau, Jeremy & Weiss, Michael (eds.), *Multi Nominis Grammaticus. Studies in Classical and Indo-European Linguistics in Honor of Alan J. Nussbaum on the Occasion of his Sixty-Fifth Birthday*, Beech Stave, Ann Arbor–New York: 175–184
- Melchert, H. Craig 2012 [2014], "Dative-locative objects of infinitives in anatolian," *Historische Sprachforschung* 125: 242–247
- Melchert, H. Craig 2013, "Ablaut Patterns in the Hittite *hi*-Conjugation," in: Jamison, Stephanie W., Melchert, H. Craig & Vine, Brent (eds.), *Proceedings of the 24th Annual UCLA Indo-European Conference*, Hempen, Bremen: 137–150
- Melchert, H. Craig 2014, "Anatolian Nominal Stems in \*-(C)o-," in: Oettinger, Norbert & Steer, Thomas (eds.), *Das Nomen im Indogermanischen: Morphologie, Substantiv versus Adjektiv, Kollektivum. Akten der Arbeitstagung der Indogermanischen Gesellschaft vom 14. bis 16. September 2011 in Erlangen*, Reichert, Wiesbaden: 205–214
- Melchert, H. Craig forthcoming, "The Position of Anatolian," in: Garrett, Andrew & Weiss, Michael (eds.), *Oxford Handbook of Indo-European Studies*, Oxford University Press, Oxford
- Melchert, H. Craig 2015, "The Tocharian s-Preterite," in: Malzahn, Melanie; Michaël Peyrot, Hannes Fellner & Theresa-Susanna Illés (eds.), *Tocharian Texts in Context: International Conference on Tocharian Manuscripts and Silk Road Culture Vienna, June 25th-29th, 2013*, Hempen, Bremen: 127–135
- Melchert, H. Craig & Oettinger, Norbert 2009, "Ablativ und instrumental im Hethitischen und indogermanischen. Ein Beitrag zur relativen Chronologie," *Incontri Linguistici* 32: 53–73
- Morpurgo Davies, Anna 1979, "The Luwian languages and the Hittite *-hi* conjugation," in: Brogyanyi, Bela (ed.), *Studies in Diachronic, Synchronic and Typological Linguistics. Festschrift for Oswald Szemerényi on the Occasion of his 65th Birthday*, Benjamins, Amsterdam: 577–610
- Morpurgo Davies, Anna 1982/1983, "Dentals, rhotacism and verbal endings in the Luwian languages," *Zeitschrift für Vergleichende Sprachforschung* 96: 245–270
- Neu, Erich 1968, *Das hethitische Mediopassiv und seine indogermanischen Grundlagen*, Harrassowitz, Wiesbaden
- Nollé, Johannes 2001, *Side im Altertum. Geschichte und Zeugnisse*, Band 2, Habelt, Bonn
- Oettinger, Norbert 1978, "Die Gliederung des anatolischen Sprachgebiets," *Zeitschrift für Vergleichende Sprachforschung* 92: 74–92
- Oettinger, Norbert 1979, *Die Stammbildung des hethitischen Verbums*, Carl, Nürnberg
- Oettinger, Norbert 1986, "*Indo-Hittite* -Hypothese und Wortbildung," Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck
- Oettinger, Norbert 1987, "Bemerkungen zur anatolischen *i*-Motion und Genusfrage," *Zeitschrift für Vergleichende Sprachforschung* 100: 35–43
- Oettinger, Norbert 2001, "Varia Hethitica," *Historische Sprachforschung* 114: 80–89
- Oettinger, Norbert 2002, "Revidierte Fassung der Einteilung der hethitischen Verbalklassen," in: *Stammbildung des hethitischen Verbums*, Nachdruck, Verlag der TU Dresden, Dresden: xix–xxx

- Oettinger, Norbert 2006, "Review of Jasanoff 2003," *Kratylos* 51: 34–45
- Oettinger, Norbert 2007, "Der hethitische Imperative auf *-i* vom Typ *paḫši* 'schütze'," in: Groddek, Detlev & Zorman, Marina (eds.), *Tabularia Hethaeorum: Hethitische Beiträge Silvin Košak zum 65. Geburtstag*, Harrassowitz, Wiesbaden: 561–568
- Oettinger, Norbert 2013–2014, "Die Indo-Hittite-Hypothese aus heutiger Sicht," *Münchener Studien zur Sprachwissenschaft* 67/2: 149–176
- Oshiro, Terumasa 1993, "Notes on Hieroglyphic Luwian," *Oriente* 29: 45–56
- Pedersen, Holger 1938, *Hittitisch und die anderen indo-europäischen Sprachen*, Munksgaard, Copenhagen
- Pérez Orozco, Santiago 2007, "La lengua sidética. Ensayo de síntesis," *Kadmos* 46: 125–142
- Popko, Maciej 2008, *Völker und Sprachen Altanatoliens* [tr. Cyril Brosch], Harrassowitz, Wiesbaden
- Probert, Philomen 2006, "Clause boundaries in Old Hittite relative sentences," *Transactions of the Philological Society* 104: 17–83
- Puhvel, Jaan 1984–, *Hittite Etymological Dictionary*, Mouton-de Gruyter, Berlin
- Puhvel, Jaan 1994, "West-Indo-European affinities of anatolian," in: Dunkel, George E.; Gisela Meyer, Salvatore Scarlata & Christian Seidl (eds.), *Früh-, Mittel-Spätindogermanisch. Akten der IX. Fachtagung der Indogermanischen Gesellschaft vom 5. bis 9. Oktober 1992 in Zürich*, Reichert, Wiesbaden: 315–324
- Rasmussen, Jens 1992, "The Distribution of *e* and *a* in Lycian," in: Beekes, Robert; Lubotsky, Alexander & Weitenberg, Jos (eds.), *Rekonstruktion und Relative Chronologie. Akten der VIII. Fachtagung der Indogermanischen Gesellschaft Leiden, 31. August – 4. September 1987*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck: 359–366
- Rieken, Elisabeth 1999a, *Untersuchungen zur nominalen Stammbildung des Hethitischen*, Harrassowitz, Wiesbaden
- Rieken, Elisabeth 1999b, "Zur Verwendung der Konjunktion *ta* in den heth. Texten," *Münchener Studien zur Sprachwissenschaft* 59: 63–88
- Rieken, Elisabeth 2004, "Das Präteritum des Medio-Passivs im Hieroglyphen-Luwischen," *Historische Sprachforschung* 117: 179–188
- Rieken, Elisabeth 2005a, "Hethitisch," in: Streck, Michael P. (ed.), *Sprachen des Alten Orients*, 3. überarbeitete Auflage, Wissenschaftliche Buchgesellschaft, Darmstadt: 80–127
- Rieken, Elisabeth 2005b, "Neues zum Ursprung der anatolischen *i*-Mutation," *Historische Sprachforschung* 118: 48–74
- Rieken, Elisabeth 2008, "The Origin of the *-l* Genitive and the History of the Stems in *-il-* and *-ül-* in Hittite," in: Jones-Bley, Karlene; Martin E. Huld, Angela Della Volpe & Miriam Robbins Dexter (eds.), *Proceedings of the 19th UCLA Indo-European Conference. Los Angeles, November 3–4, 2007*, Institute for the Study of Man, Washington, DC: 239–256
- Rieken, Elisabeth 2009, "Der Archaismus des Hethitischen – eine Bestandsaufnahme," *Incontri Linguistici* 32: 37–52
- Rieken, Elisabeth 2011, "Verberststellung in hethitischen Übersetzungstexten," in: Krisch, Thomas & Lindner, Thomas (eds.), *Indogermanistik und Linguistik im Dialog: Akten der XIII. Fachtagung der Indogermanischen Gesellschaft vom 21. bis 27. September 2008 in Salzburg*, Reichert, Wiesbaden: 498–507
- Rizza, Alfredo 2007, *I pronomi enclitici nei testi etei di traduzione dal hattico*, Iuculano, Pavia

- Shatskov, Andrej 2006, "Some remarks on Hittite infix verbs of the type *harni(n)k-*," *Altorientalische Forschungen* 33: 286–292
- Sideltsev, Andrej V. *forthcoming*, "Clause internal and clause leftmost verbs in Hittite"
- Starke, Frank 1977, *Die Funktionen der dimensionalen Kasus und Adverbien im Althethitischen*, Harrassowitz, Wiesbaden
- Starke, Frank 1985, *Die keilschrift-luwischen Texte in Umschrift*, Harrassowitz, Wiesbaden
- Starke, Frank 1990, *Untersuchung zur Stammbildung des keilschrift-luwischen Nomens*, Harrassowitz, Wiesbaden
- Strunk, Klaus 1994, "Relative chronology and Indo-European verbal-system: The case of present- and aorist-stems," *Journal of Indo-European Studies* 22: 417–434
- Sturtevant, Edgar H. 1933, *A Comparative Grammar of the Hittite Language*, Linguistic Society of America, Philadelphia
- Tischler, Johann 1977–2016, *Hethitisches Etymologisches Glossar*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck
- Tischler, Johann 1979, "Der indogermanische Anteil am Wortschatz des Hethitischen," in: Neu, Erich & Meid, Wolfgang (eds.), *Hethitisch und Indogermanisch*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck: 257–267
- Tischler, Johann 2001, *Hethitisches Handwörterbuch*, Institut für Sprachen und Literaturen der Universität Innsbruck, Innsbruck
- Watkins, Calvert 1963, "Preliminaries to a historical and comparative analysis of the syntax of the Old Irish verb," *Celtic* 6: 1–49
- Watkins, Calvert 1969, *Geschichte der indogermanischen Verbalflexion. Indogermanische Grammatik*, ed. by Jerzy Kuryłowicz, Band III: Formenlehre, Erster Teil, Winter, Heidelberg
- Weeden, Mark 2011, Review of Kloekhorst 2008, *Bulletin of the School of African and Oriental Studies* 74/1: 59–76
- Weitenberg, Joseph J.S. 1984, *Die hethitischen U-Stämme*, Rodopi, Amsterdam
- Weitenberg, Joseph J.S. 1992, "The use of asyndesis and particles in Old Hittite simple sentences," in: Carruba, Onofrio (ed.), *Per una Grammatica hittita*, Iuculano, Pavia: 305–353
- Woodard, Roger (ed.) 2004, *The Cambridge Encyclopedia of the World's Ancient Languages*, Cambridge University Press, Cambridge
- Yakubovich, Ilya 2010, *Sociolinguistics of the Luvian Language*, Brill, Leiden–Boston
- Yoshida, Kazuhiko 1990, *The Hittite Mediopassive Endings in -ri*, de Gruyter, Berlin–New York
- Yoshida, Kazuhiko 1993, "Notes on the prehistory of Preterite verbal endings in Anatolian," *Historische Sprachforschung* 106: 26–35
- Yoshida, Kazuhiko 2010, "1st singular iterated mediopassive endings in Anatolian," in: Jamison, Stephanie W.; Melchert, H. Craig & Vine, Brent (eds.), *Proceedings of the 21st Annual UCLA Indo-European Conference*, Hempfen, Bremen: 231–243



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

PART 4

# **INDO-IRANIAN**





Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

# INDO-IRANIAN

*Leonid Kulikov*

## INDO-IRANIANS AND THEIR ORIGINS

Indo-Iranian languages (another, outdated and quite misleading term is “Aryan”), which form a major branch of the Indo-European language family, are spoken by more than a billion of speakers occupying an immense territory from the Caucasus and southeastern Anatolia in the West to Northeast India and Bengal in the East and the Maldivic Islands in the South. This grouping includes the Indo-Aryan, Iranian, Nūristānī (in earlier scholarship also called Kāfir), and Dardic branches. The latter is considered by many scholars as part of the Indo-Aryan branch (e.g., Morgenstierne 1961, 1973, Strand 1973, Southworth 2005: 127, 149). Convincing argumentation for treating the Dardic languages as a separate genetic group, a fourth sub-branch of Indo-Iranian, alongside Indo-Aryan, Iranian, and Nūristānī, can be found in Kogan 2005.

All Indo-Iranian languages derive from a common ancestor, the reconstructed Proto-Indo-Iranian language that was spoken around the end of the 3rd millennium BCE. The speakers of this proto-language, Proto-Indo-Iranians, are commonly identified with the Sintashta archaeological culture and its successor, the (early) Andronovo culture, located in the vast area between the Southern Ural and the Tian Shan mountains and dated to the end of the 3rd millennium BCE to the 1st half of the 2nd millennium BCE.

The oldest preserved linguistic evidence comes from the Mitanni kingdom, which existed between ca. 1500 BCE and 1300 BCE on the territory that is now northern Syria and southeastern Anatolia (and thus almost precisely matching with the easternmost end of the territory populated by the speakers of the modern Indo-Iranian languages). The main language of the kingdom was Hurrian (belonging to the Hurro-Urartian language family, probably related to the Northeast Caucasian, or Nakh-Daghestanian, family). However, the names of the ruling elite (such as Artaššumara [~ Ilr. *Arta-smara* ‘who thinks of Arta/Ṛta’] or Biridashva (biridašva, biriiša) [~ *Prītāšva* ‘whose horse is dear’]), and, particularly, the names of the gods (Mitra, Varuṇa, Indra), as well as a number of technical terms preserved in a Hittite treatise on horse training (such as *aika* ‘one’ [cf. Ved. Skr. *éka*], *panza* (~ *pāñca*) ‘five’), *vartana* ‘round’), preserve traces of an Indo-Iranian language (or perhaps, more precisely, Indo-Aryan, as a few features may point to an Indo-Aryan, rather than common Indo-Iranian, character for this hypothetical idiom; see Mayrhofer 1974).

The actual text documentation starts from the 2nd half of the 2nd millennium BCE (ca. 1200–1000 BCE) for Indo-Aryan, with the hymns of the Ṛgveda (RV), the main sacred text of the Hinduism, while the text documentation of Iranian begins in the 1st half of the 1st millennium BCE (Avesta).

Due to the essentially uncontroversial character of the Proto-Indo-Iranian reconstruction, Indo-Iranian remains probably the best-established larger subgrouping within the Indo-European family and the only one that is adopted by all Indo-Europeanists (even another very likely subgrouping of approximately the same level, Balto-Slavic, is doubted by a number of scholars).

## INDO-IRANIAN PHONOLOGY

The most important features characterizing Proto-Indo-Iranian as opposed to all or most other branches of Indo-European include:

- (i) The merger of the two PIE vowels, *e* and *o*, in *a*, with the preservation of the quantity distinctions, i.e., PIE \**e*, \**o* > PIIr. \**a*; PIE \**ē*, \**ō* > PIIr. \**ā*; cf. Ved. *dāśa*, Av. *dasā* ‘ten’ (~ Gr. δέκα, Lat. *decem*) < PIE \**dek̑m*; Ved. *vāk*, Av. *vāxš* ‘speech’ (~ Lat. *uox* ‘voice’) < PIE \**wōk*<sup>ws</sup>. The same holds for the vocalic part of the PIE diphthongs, which invariably changes to *a*, i.e., PIE \**ey*, \**oy* > PIIr. \**ai*; etc. The only phonological context where the original opposition between two PIE timbres leaves traces in the form of a new opposition (in length) is the open non-final syllable, where the original short \**o* was lengthened (\**o* | \_CV > \**ā*), while \**e* was not (Brugmann’s Law); cf. the acc. sg. form Ved. *dātāram* ‘giver’, Av. *dātārəm* ‘creator’ (~ Gr. δώτορα) < PIE \**-tor-ṃ* as opposed to acc. sg. Ved. *pitāram*, Av. *pitārəm* ‘father’ (~ Gr. πατέρα) < PIE \**-ter-ṃ*.
- (ii) Another source of PIIr. \**a* is the vocalic allophone of the two nasal sonants, i.e., PIE \**ṅ*, \**ṃ* > PIIr. \**a*; cf. Ved. *matī-* ‘thought’ (~ Lat. *mēns*, gen. sg. *mentis* ‘mind’) < PIE \**mn-ti-* [\**mṃ-ti-*] and the words for ‘ten’ mentioned above.
- (iii) The phonological status of \**l* was weakened in Indo-Iranian. *l* changed to *r* in Iranian, and in many, but not all, cases in Indo-Aryan. In particular, *l* always changed to *r* when surfacing as a vocalic allophone; in the language of the RV, words containing *l* are very rare, but from the end of the early Vedic period, more forms with *l* start to appear (perhaps originating in one or several hypothetical *l*-dialects, where the original opposition was preserved); cf. RV *reh-* as opposed to the younger Atharvavedic (Paippalāda) root variant *leh-* ‘lick’ < PIE \**leyǵ*<sup>h</sup>.
- (iv) Proto-Indo-Iranian is very conservative as far as the system of stop consonants is concerned, preserving the PIE system virtually intact: the opposition of three series – voiceless T, voiced non-aspirate D, and voiced aspirate D<sup>h</sup> (or glottalized ’D and simple voiced D, if we accept the glottalic theory for the PIE reconstruction) – remains without changes. Voiceless aspirates (and, within the glottalic theory, voiced aspirates, too) must represent Indo-Aryan innovation.
- (v) Palatalizations: Indo-Iranian belongs to the *satəm* languages; that is, the simple velar (\**k* \**g* \**g*<sup>h</sup>) and labiovelar (\**k*<sup>w</sup> \**g*<sup>w</sup> \**g*<sup>w</sup><sup>h</sup>) PIE series of consonants merged together, while the palatovelars were palatalized: \**k* \**ǵ* \**ǵ*<sup>h</sup> > \**č* \**j* \**j*<sup>h</sup>; cf. Ved. *dāśa*, Av. *dasā* ‘ten’ < PIIr. \**daśa* < PIE \**dek̑m* (~ Gr. δέκα etc.). Another series of palatal consonants, PIIr. \**č* \**j* \**j*<sup>h</sup>, reflects PIE simple velars and labiovelars before /*ē*/ and /*i*/; cf. Ved. *ca*, Av. *-ča* ‘and’ < PIE \**k<sup>w</sup>e* (~ Lat. *que* etc.).
- (vi) The only PIE sibilant, \**s*, is retroflexivized in PIIr. after *i*, *u*, *r*, and *K* (= all velars); cf. Ved. *tṛṣṇā-*, Av. *taršna-* ‘thirst’ < PIE \**tṛs-n-* (~ Eng. *thirsty*); Ved. *viśā-*, Av. *vīša-* ‘poison’ < PIE \**wis-ó-* (~ Lat. *vīrus*). This phonetic law, known also as the RUKI rule, is an isogloss (partly) shared with the Balto-Slavic, Armenian, and Albanian branches.
- (vii) Indo-Iranian differs from all other branches in the timbre of the vocalization of the PIE laryngeals (in the interconsonantal position): \**Ḥ* > \**i* (unconditionally in the final syllable and occasionally also in other positions); cf. PIE \**ph<sub>2</sub>-ter-* ‘father’ > PIIr. \**p(i)tar-* > Ved. *pitár-*, Av. dat. sg. *piḍrē, fəḍrōi*, OPers. *pitár-*.
- (viii) Indo-Iranian is the most conservative branch of Indo-European as far as the accentuation system is concerned: Vedic is the main source of our knowledge about PIE accentuation, preserving it, presumably, virtually intact.

## THE INDO-IRANIAN MORPHOLOGICAL SYSTEM

From the morphological point of view, Indo-Iranian is also one of the most archaic branches of Indo-European, probably preserving the original system of inflectional and derivational categories as reconstructable for the proto-language better than any other branch.

In particular, within the nominal system Indo-Iranian perfectly preserves the three numbers (singular, dual, and plural) and eight cases (nominative, accusative, instrumental, dative, ablative, genitive, locative, and vocative) as well as the three genders of PIE. Only some regularizations in the system of endings in nominal paradigms must represent Indo-Iranian innovations, though. These include, for instance, introducing *y* after *b<sup>h</sup>* in the dative-ablative dual and plural endings \*-b<sup>h</sup>yā(m) and \*-b<sup>h</sup>yas, respectively (for \*-b<sup>h</sup>ō and \*-b<sup>h</sup>os reconstructable on the basis of evidence from other branches), apparently from the instrumental plural \*-b<sup>h</sup>i-s.

Another probable innovation in the Indo-Iranian pronominal system was adding (prefixing or suffixing) the particle of remoteness \*u/\*au to certain forms derived from the pronominal stem \*a- (cf. Iranian \*au-a- and Indo-Aryan \*ad-u → \*ad-áu by analogy with \*sáu; \*am-ú-), in order to distinguish between proximal and distal demonstratives, presumably due to the fact that some originally distinct pronouns fell together after the merger of the two PIE vowels \*e and \*o in Indo-Iranian (Klein 1977).

The Indo-Iranian verbal system was also quite archaic. It had five moods: indicative, injunctive (very productive in the earliest periods, where it was used as an “extratemporal” tense category rather than a special mood” (M. Kümmel *forthcoming*/2017), mentioning an event or activity as a “known fact”), imperative, optative, and subjunctive; and it distinguished between the present tense and three past tenses (aorist, imperfect, and perfect) in the indicative. The function of the future was largely taken over by the subjunctive mood, which denoted the expected action (cf. Tichy 2006).

Indo-Iranian had two voices, active and middle; at least in the 3 sg. and 3 pl. forms, many verbs could distinguish between “(medio-)passive” and “non-passive”. The “(medio-)passive” paradigm was differentiated from the non-passive forms partly by means of special stem formative morphemes (such as the present passive suffix \*-ja-, the use of which in this function was an important Indo-Iranian innovation), partly by means of specific inflection (see Kümmel 1996, Kulikov 2006).

## THE INDO-IRANIAN LEXICON

### The inherited IE lexicon

A large part of the Common Indo-Iranian vocabulary is inherited from PIE, being, at the same time, an important source for its reconstruction. This inherited part of the lexicon covers all important domains of lexica (see numerous examples above and in the Indo-Aryan and Iranian chapters) and, like other parts of the linguistic system, testifies to a considerable conservatism of Indo-Iranian.

### Loan-words in the Indo-Iranian lexicon and substrate/adstrate influence

Alongside with the well-established inherited part of the lexicon, there are a considerable number of words that cannot be traced to PIE but are safely reconstructable for Proto-Indo-Iranian on the basis of evidence from both Old Indo-Aryan and (Old) Iranian (evidence from Middle and New Iranian may be relevant, as the expected cognates may

be lacking from the Old Iranian corpus, which is somewhat more limited as compared to Old Indo-Aryan), thus representing loan-words into Proto-Indo-Iranian before its split. Such borrowings encompass a variety of cultural domains, such as agricultural and farming terms, animal names (cf. Ved. *úštra-*, Av. *uštra-*, OPers. *uša-* < PIIr. \**úštra-* ‘camel’; Ved. *kaśyápa-*, Av. *kasiapa-* < PIIr. \**kačjápa-* ‘tortoise’), household and building (cf. Ved. *ištakā-*, *ištikā-*, Av. *ištīia-*, OPers. *išti-* < PIIr. \**išta-*, \**išti-* ‘brick’), irrigation terms, religious terms, and mythology (Ved. *aṁśú-*, Av. *qsu-* < PIIr. \**ańcú-* ‘Soma plant’; Ved. *átharvan-*, Av. *āθrauuān-* < PIIr. \**át<sup>h</sup>arūan-* ‘(a particular type of) priest’ (? see Pinault 2006)).

Candidates for the sources of these borrowings may be cultural languages that were spoken in and around the area populated by the Proto-Indo-Iranians, that is, from Mesopotamia to the South Ural, around the 3rd millennium BCE, as well as some isolated languages spoken in remote mountain areas, such as Burushaski (see Tikkanen 1988). Proto-Uralic could also borrow loan-words from some Indo-Iranian dialects; see, e.g., Katz 2003. However, the origin of most of these loan-words is unclear. Quite attractive is the theory about the borrowing of the bulk of these Common Indo-Iranian loan-words from the unidentified language spoken by the population of the Bactria-Margiana Archaeological Complex that was located immediately to the south of the Andronovo culture’s area and had a well-developed urban civilization (see Lubotsky 2001).

## INDO-IRANIAN LANGUAGES IN A DIACHRONIC TYPOLOGICAL PERSPECTIVE

The relevant shared features of the languages belonging to the Indo-Iranian branch are not confined to the peculiarities inherited by Proto-Indo-Iranian from PIE and thus are shared, foremost, by the oldest representatives of this subgrouping, i.e., by Old Indo-Aryan (Vedic/Sanskrit) and Old Iranian (Avestan and Old Persian), but they are only partly (if at all) preserved in the modern members of the branch. Of no less importance are much more recent features of the group that emerged after the split of Proto-Indo-Iranian (or even of Proto-Indo-Aryan and Proto-Iranian) and must be due to the fact that many members of the Indo-Iranian branch belong to the same linguistic area (together with languages of non-Indo-European affiliation). Specifically, virtually all Indo-Aryan and Nūristānī languages as well as a few neighboring Iranian languages such as Pashto (together with the non-Indo-European Dravidian, Munda, and the isolate Burushaski) form the South Asian linguistic area (which is notorious as one of the handbook examples of a linguistic area; see Masica 1976). There are also some good reasons to extend this area further to the North and East, including such languages or families as Tocharian (Indo-European) and Turkic/Altaic into the large Central-South Asian linguistic area (see, e.g., Hock 2007).

The features briefly summarized below thus reveal the same basic type of evolution, representing shared innovations due to parallel developments and/or triggered by shared influences of neighboring (substrate or adstrate) languages. The list of linguistic features given in this section by no means claims to be exhaustive.

### Diachronic trends within the system of transitivity oppositions: voice and valency-changing categories

Considering the system of transitivity oppositions in Indo-European in a diachronic typological perspective, one may posit two major evolutionary types attested for the

members of this linguistic family (briefly outlined in Comrie 2006; for details, see Kulikov 2009; 2012). One major type is located in the West of the Indo-European area and includes such branches as Germanic, Romance, Greek, and Slavic. These branches share a number of parallel tendencies in the system of encoding transitivity oppositions and valency-changing categories that can be qualified, by and large, as instantiating a “**syncretic**” type of evolution: (i) The old syncretic marker encoding a number of valency-reducing (intransitivizing) derivations (such as passive, anticausative, and reflexive), the middle voice (type of inflection), is either preserved (as in Greek) or replaced with a new middle, mostly going back to the PIE reflexive pronoun \*s(w)e-. (ii) The PIE morphological causatives with the suffix \*-éye/o-, reflexes of which are still relatively well attested in Old Germanic (e.g., in Gothic *jan*-verbs) and (early) Slavic (*i*-causatives), virtually disappear. (iii) Many languages of this “Western” area attest the emergence and expansion of labile verbs (i.e., verbs or verbal forms that can show changes in syntactic pattern, with no formal change in the verb) of the type *John opens the door* / *The door opens* (this trend is particularly well attested in the history of English; see, e.g., McWhorter 2002, Gelderen 2011).

By contrast, several other daughter languages, mostly those belonging to some **Eastern** branches of Indo-European, such as, first of all, Indo-Iranian, radically abandon the syncretic strategy in the system of encoding transitivity oppositions and valency-changing categories, developing, by and large, in the opposite, “**antisyncretic**”, direction. The most important diachronic typological features shared by the representatives of the Eastern type include the following: (i) The middle voice declines, which amounts to the loss of most original (intransitivizing) functions of the PIE middle and the lexicalization of many middle forms; subsequently, the active/middle opposition, albeit physically preserved in the paradigm, loses a large part of its functional content, thus being degrammaticalized. (ii) Instead, most languages of this area develop special markers for several intransitivizing voices and valency-changing categories. These include, in particular, Indo-Iranian and Armenian markers of a morphological passive going back to the PIE suffix \*-ye/o- and Indo-Iranian reflexive and reciprocal pronouns (cf. Ved. *tanū́-*, Av. *tanū-* ‘self’ ← ‘body’; Ved. *anyó . . . anyá-* (*anyo-* ‘*nya-*, *anyonya-*), Av. *aniiō. ainīm* ‘one another, each other’; see Kulikov 2014b). Morphological causatives (in particular, the reflexes of the PIE \*-éye/o-causative – cf. Vedic and Avestan morphological causatives with the suffixes *-áya-* and *-aiia-*, respectively – but also other morphological causatives, especially with nasal affixes) demonstrate a considerable increase in productivity. This increase in the productivity of morphological causatives can be observed during the entire history of Indo-Iranian languages and eventually gives rise to the complex systems with two or even more causatives (including double causatives with a reiterated causative morpheme, as in Marathi and Tajik; see examples in Masica 1976: Ch. 3) in some members of this branch, which distinguishes them from other Indo-European groups. (iii) Labile verbs, still common in (early) Old Indo-Aryan and Avestan, become rarer; in particular, in Vedic this tendency is clearly observable as early as in the post-Ṛgvedic language; see Kulikov 2014a.

Typologically, the antisyncretic (Eastern) type, as attested in Indo-Iranian, shares more features with the languages of some non-Indo-European families, such as Turkic (or, more generally, Altaic) or Dravidian, than with the languages of several Western Indo-European branches. This is demonstrated in Table 4.1, which compares a few related features characterizing the main trends in the development of the systems of transitivity oppositions for a variety of both Indo-European and non-Indo-European language groups.

**TABLE 4.1 ENCODING TRANSITIVITY OPPOSITIONS: DIACHRONIC TYPOLOGICAL FEATURES OF SOME LANGUAGE FAMILIES**

	middle voice	morphological valency-changing categories	labile syntax
East Caucasian	0	+/ $\nearrow$	+
Kartvelian	+	+	0
Egyptian/Coptic	+	+/ $\nearrow$	(+)
Indo-European			
Germanic	$\nearrow$ (new)	$\searrow$	$\nearrow$
Romance	$\nearrow$ (new)	$\searrow$	+/ $\nearrow$
Slavic	$\nearrow$ (new)	$\searrow$	0/(+)
Greek	+	0/(+)	++
Indo-Iranian	$\searrow$	$\nearrow$	$\searrow$
Turkic	0/ $\searrow$	+/ $\nearrow$	0/ $\searrow$

0	lacking
(+)	present but rare or non-productive
+	present
$\nearrow$	increasing (in productivity, frequency, etc.)
$\searrow$	decreasing/disappearing

### Diachronic changes within the case system

As mentioned above, the Proto-Indo-Iranian morphological system was very conservative, perfectly preserving the eight PIE cases, which further remained intact in both main branches, Indo-Aryan and Iranian, for more than 1000 years after the split of Proto-Indo-Iranian, as documented in Old Indo-Aryan (Vedic) and Old Iranian (Avestan). After the end of the Old Indo-Aryan and Old Iranian periods, we observe a dramatic decline of the case systems in both subgroups, which thus represents another shared innovation. By the end of the Middle Indo-Aryan/Iranian period, the Indo-Iranian languages had lost the majority of the conservative Old Indo-Aryan/Iranian case system. Generally, only two cases survive, direct (resulting from the merger of nominative and accusative) and oblique (mostly going back to the Old Indo-Aryan/Iranian genitive); in some languages traces of some other oblique cases, such as the instrumental, locative, or ablative, can still be found, such as the Sinhala instrumental case in *-en/-in* and the Assamese ergative *-e*, both probably reflecting the OIA instrumental singular ending of the *-a*-declension, *-ena*. The functions of the lost cases are largely taken over by postpositions of different origin.

In several New Indo-Aryan and New Iranian languages we observe the grammaticalization of such new postpositions, which are normally added to the oblique case form. Very often this grammaticalization results in the amalgamation of a postposition with the oblique and, ultimately, in the rise of a new case. Such is, for instance, the origin of the two new case endings in Sinhala, viz. dat. *-ta* (< Skt. *artha* ‘goal, purpose’, used as the final member of a nominal compound) and gen. *-ge* (< Skr. *gr̥he* ‘in the house’ = loc. sg. of *gr̥ha-* ‘house’), and several case morphemes in Ossetic, such as dat. *-æn* (< \**ana*, cf. Av. *ana* ‘upon, over, across’, or \**anu*, cf. OPers. *anuv*, Av. *anu* ‘along, after, according to’) and adessive *-bæl* (Digoron) (< \**upari* ‘above, upon, on’); for details and discussion, see Kulikov 2011.

### Other shared syntactic innovations

There are a number of other syntactic innovations shared by many or most Indo-Iranian languages of both major subgroups:

- (i) The emergence of the ergative construction, first of all, on the basis of the syntactic reanalysis of constructions with resultative verbal adjectives (traditionally

also called “perfect passive participles”) with the suffix *-ta-/-na-*, is a syntactic innovation shared by several Indo-Aryan and Iranian (foremost Western Iranian) languages. This development started during the Middle Indo-Aryan/Iranian period, when the agentive noun (the passive agent) of constructions with *-ta-/-na-* adjectives acquired some subject properties (such as the referential control of converbs). In fact, however, the origins of the ergative alignment can presumably be dated as early as Old Indo-Aryan/Iranian or perhaps Proto-Indo-Iranian. In particular, there are some reasons to take Vedic constructions with *-ta-/-na-* adjectives and the genitive marking of the agent separately from canonical passives with the instrumental. According to P. K. Andersen (1986), the genitive noun displays a number of subject properties (usually animate; definite and/or refers to old information) in such constructions, and therefore they should be qualified as ergative rather than passive, properly speaking.

- (ii) Constructions with oblique (non-canonical) subjects, quite common in many Indo-European branches (in particular, in the Western part of the Indo-European area; see, e.g., Barðdal et al. 2012), had become relatively rare in the easternmost branches (in particular, in Indo-Iranian), being replaced by the canonical (nominative) subject pattern. This innovation appears to be shared with some other Eastern (antisyncretic) branches of Indo-European, such as Tocharian; see Viti (forthcoming). Nevertheless, many New Indo-Aryan languages have ultimately developed new oblique (non-canonical) subject constructions (see Masica 1991: 346ff.).
- (iii) A remarkable feature of the syntax of many modern Indo-Iranian languages is the existence of long chains of converbs in narrative texts (see, e.g., Masica 1976). This feature is shared with a number of non-Indo-European languages of South, East, and Central Asia (cf. Bickel’s (1998) “Asian type”).

The origin of the cluster of (diachronic) typological features described above, such as, above all, the antisyncretic evolutionary type of encoding transitivity oppositions, the rise of new agglutinating case systems, and the decline of constructions with non-canonical (oblique) subjects in Old Indo-Aryan, is a difficult problem on its own. It may be partly due to the influence of substrate (or adstrate) languages of the Altaic or Dravidian type. Further study of these areal features in typological context is required for a better understanding of the mechanisms of these changes in the historical period.

## FURTHER READING

There is no comprehensive monographic description of Proto-Indo-Iranian and its historical grammar. However, the relevant chapters of de Gruyter’s forthcoming *Handbook of Comparative and Historical Indo-European Linguistics*, Lubotsky forthcoming/2017, Kümmel forthcoming/2017, and Schmitt forthcoming/2017, give excellent brief surveys of the corresponding parts (phonology, morphology, and the lexicon, respectively) of the Proto-Indo-Iranian linguistic system. In addition, Gotō 2013 offers a comprehensive overview of the Indo-Aryan morphological system in a diachronic perspective, tracing most of its aspects back to the Indo-Iranian. The reconstructable Indo-Iranian vocabulary is fairly well represented in the Indo-Aryan etymological dictionary Mayrhofer 1992–1996.

There are several works on the South Asian linguistic area that discuss the features shared by its members, i.e., virtually all Indo-Aryan, Nūristānī, and some Iranian languages, starting with such seminal works as Masica 1976 (see also Masica 1991); see also, more recently, Southworth 2005 and Hock 2007.



## REFERENCES

- Andersen, Paul Kent 1986, "The genitive agent in Rigvedic passive constructions," in: Slawski, F. (ed.), *Collectanea linguistica in honorem Adami Heinz*, Zakład Narodowy im. Ossolińskich, Wrocław: 9–13
- Barðdal, Jóhanna et al. 2012, "Reconstructing constructional semantics: The dative subject construction in Old Norse-Icelandic, Latin, Ancient Greek, Old Russian and Old Lithuanian," *Studies in Language* 36/3: 511–547
- Bickel, Balthasar 1998, "Converbs in cross-linguistic perspective" [rev. article of Haspelmath, M. & König, E. (eds.), *Converbs*, Berlin: Mouton de Gruyter, 1995], *Linguistic Typology* 2/3: 381–397
- Comrie, Bernard 2006, "Transitivity pairs, markedness, and diachronic stability," *Linguistics* 44/2: 303–318
- Gelderen, Elly van 2011, "Valency changes in the history of English," *Journal of Historical Linguistics* 1/1: 106–143
- Gotō, Toshifumi 2013, *Old Indo-Aryan Morphology and its Indo-Iranian Background* / in co-operation with J.S. Klein and V. Sadovski, Wien
- Hock, Hans H. 2007, "South Asia and Turcic: The Central Asian connection?," in: Masica, C.P. (ed.), *Old and New Perspectives on South Asian Languages: Grammar and Semantics*, Motilal Banarsidass, Delhi: 65–90
- Katz, Hartmut 2003, *Studien zu den älteren indoiranischen Lehnwörtern in den uralischen Sprachen*, Winter, Heidelberg
- Klein, Jared S. 1977, "The Indo-Iranian prehistory of the Sanskrit *asáu/amám* pronoun," *Journal of Indo-European Studies* 5: 161–176
- Kogan, Anton I. 2005, *Dardskie jazyki: genetičeskaja xarakteristika* [Dardic languages: Genetic characterization], Vostočnaja Literatura, Moskva
- Kulikov, Leonid 2006, "Passive and middle in Indo-European: Reconstructing the early Vedic passive paradigm," in: Abraham, W. & Leisiö, L. (eds.), *Passivization and Typology: Form and Function*, Benjamins, Amsterdam: 62–81
- Kulikov, Leonid 2009, "Valency-changing categories in Indo-Aryan and Indo-European: A diachronic typological portrait of Vedic Sanskrit," in: Saxena, A. & Viberg, Å. (eds.), *Multilingualism. Proceedings of the 23rd Scandinavian Conference of Linguistics*, Uppsala University, 1–3 October 2008 (*Studia Linguistica Upsaliensia*; 8), Uppsala Universitet, Uppsala: 75–92
- Kulikov, Leonid 2011 [2012], "The Proto-Indo-European case system and its reflexes in a diachronic typological perspective: Evidence for the linguistic prehistory of Eurasia," *Rivista degli studi orientali* 84: 289–309
- Kulikov, Leonid 2012, "Voice and valency derivations in Old Indo-Aryan in a diachronic typological perspective: The degrammaticalization of the middle and other trends in the Vedic verbal system," in: Melazzo, Lucio (ed.), *Usare il presente per spiegare il passato. Teorie linguistiche contemporanee e lingue storiche. Atti del XXXIII Convegno della Società Italiana di Glottologia*, Palermo, 16–18 ottobre 2008 (Biblioteca della Società Italiana di Glottologia; 35), Il Calamo, Roma: 161–191
- Kulikov, Leonid 2014a, "The decline of labile syntax in Old Indo-Aryan: A diachronic typological perspective," *Linguistics* 52/4 (special issue: *Typology of labile verbs: Focus on diachrony*): 1139–1165
- Kulikov, Leonid 2014b, "Grammaticalization of reciprocal pronouns in Indo-Aryan: Evidence from Sanskrit and Indo-European for a diachronic typology of reciprocal constructions," *Journal of South Asian Languages and Linguistics* 1/2: 117–156

- Kümmel, Martin 1996, *Stativ und Passivaorist im Indoiranischen*, Vandenhoeck & Ruprecht, Göttingen
- Kümmel, Martin (forthcoming/2017), “The morphology of Indo-Iranian,” in: Klein, Jared; Joseph, Brian & Fritz, Matthias (eds.), *Handbook of Comparative and Historical Indo-European Linguistics. (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK); 41)*, Walter de Gruyter, Berlin
- Lubotsky, Alexander 2001, “The Indo-Iranian substratum,” in: Carpelan, C., Parpola, A. & Koskikallio, P. (eds.), *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations. (Mémoires de la Société Finno-ougrienne; 242)*, Suomalais-ugrilainen Seura, Helsinki: 301–317
- Lubotsky, Alexander (forthcoming/2017), “Proto-Indo-Iranian phonology,” in: Klein, Jared; Joseph, Brian & Fritz, Matthias (eds.), *Handbook of Comparative and Historical Indo-European Linguistics. (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK); 41)*, Walter de Gruyter, Berlin
- Masica, Colin P. 1976, *Defining a Linguistic Area: South Asia*, University of Chicago Press, Chicago
- Masica, Colin P. 1991, *The Indo-Aryan Languages (Cambridge Language Surveys)*, Cambridge University Press, Cambridge
- Mayrhofer, Manfred 1974, *Die Arier im Vorderen Orient – ein Mythos? (Sitzungsberichte der Oesterreichischen Akademie der Wissenschaften 294,3)*, Verlag der Österreichischen Akademie der Wissenschaften, Vienna
- Mayrhofer, Manfred 1992–1996, *Etymologisches Wörterbuch des Altindoarischen, I–II*, Winter, Heidelberg
- McWhorter, John H. 2002, “What happened to English?” *Diachronica* 19/2: 217–272
- Morgenstierne, Georg 1961, “Dardic and Kāfir languages,” in: Gibb, Hamilton Alexander Rosskeen et al. (eds.), *The Encyclopedia of Islam*, Vol. 2, Brill, Leiden: 138–139
- Morgenstierne, Georg 1973, *Irano-Dardica*, Reichert, Wiesbaden
- Pinault, Georges-Jean 2006, “Further links between the Indo-Iranian substratum and the BMAC language,” in: Tikkanen, B. & Hettrich, H. (eds.), *Themes and Tasks in Old and Middle Indo-Aryan Linguistics*, Motilal Banarsidass, Delhi: 167–196
- Schmitt, Rüdiger (forthcoming/2017), “The lexicon of Indo-Iranian,” in: Klein, Jared; Joseph, Brian & Fritz, Matthias (eds.), *Handbook of Comparative and Historical Indo-European Linguistics. (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK); 41)*, Walter de Gruyter, Berlin
- Southworth, Franklin C. 2005, *Linguistic Archaeology of South Asia*, Routledge Curzon, London
- Strand, Richard F. 1973, “Notes on the Nūristānī and Dardic languages,” *JAOS* 93/3: 297–305
- Tichy, Eva 2006, *Der Konjunktiv und seine Nachbarkategorien. Studien zum indogermanischen Verbum, ausgehend von der älteren vedischen Prosa*, Hempfen, Bremen
- Tikkanen, Bertil 1988, “On Burushaski and other ancient substrata in northwestern South Asia,” *Studia Orientalia* 64: 303–325
- Viti, Carlotta (2016), “The morphosyntax of experience predicates in Tocharian,” *Cahiers de Linguistique – Asie Orientale (CLAO)* 45: 26–70

# INDO-ARYAN

*Leonid Kulikov*

## INTRODUCTION

### General information

The Indo-Aryan languages (sometimes also referred to, misleadingly and not quite correctly, as Indic) represent the largest group of the Indo-European both by the total number of speakers of the present-day Indo-Aryan languages (approx. one and a half billion of the total three billion speakers of Indo-European languages) and by the number of languages (ca. 225 languages recognized, for instance, by Ethnologue, thus making up more than half of all Indo-European languages listed by this source). The largest Indo-Aryan languages include Hindi and Urdu (about 240 million speakers), Bengali (about 230 million), Punjabi (about 110 million), Marathi (about 70 million), Rajasthani (about 50 million) and Gujarati (about 45 million). At present, Indo-Aryan languages are spoken, above all, on the Indian subcontinent, also referred to as South Asia. This region includes India proper (officially the Republic of India), Bangladesh, Pakistan, Nepal, Bhutan, Sri Lanka and the Maldives (Republic of the Maldives). Due to labor migration, large groups of speakers of Indo-Aryan languages have settled all over the world, and these languages are now spoken in a number of diasporas, in particular, in such regions as South-East Asia, Africa, South America (in particular, Guyana and Surinam), Australia and Melanesia (in particular Fiji), as well as in the USA and some European countries, foremost in the United Kingdom.

The Indo-Aryan group is also the easternmost branch of the Indo-European languages (except for the extinct Tocharian languages, which were spoken still further to the East). The Indo-Aryan-speaking peoples probably came to the Indian subcontinent from the North-West, entering the greater Punjab (where the oldest Vedic text, the Ṛgveda, was composed) by the end of the first half of the second millennium BC. In that epoch, Proto-Indo-Aryan was a single language (still very close to Proto-Iranian), which started to split into dialects essentially after the arrival in India.

### Chronology of the Indo-Aryan languages and the sociolinguistic situation in ancient India

The most ancient Indo-Aryan language is Vedic (Vedic Sanskrit), characterized as the main representative of the Old Indo-Aryan (OIA) linguistic period in the development of the Indo-Aryan languages, which is attested from the second millennium BC onwards. Chronologically, Vedic can be divided into at least two main periods, early Vedic (also called “mantra language”, i.e. the language of the hymns addressed to the Vedic gods, mantras and magic spells, and middle/late Vedic (also called “Vedic prose”). The oldest layer of Vedic is attested in the language of the Ṛgveda (RV), which remains the main sacred text of the Hindus and which can approximately be dated to the second half of the second millennium BC. Within the RV, we can distinguish between the early RV (“family books”, or maṇḍalas, which include books II–VII) and the late RV (encompassing, above all, maṇḍalas I and X, as well

as a part of book VIII, Vāḷakhilya); books VIII and IX are chronologically rather heterogeneous. The language of the second most ancient text, the Atharvaveda (AV), resembles in many respects (and is essentially synchronic with) the language of the late RV. Early Vedic is followed by middle and late Vedic (= the language attested in the Brāhṃaṇas, Āraṇyakas, oldest Upaniṣads and Sūtras); the post-Vedic period covers younger Upaniṣads and Sūtras, as well as Epic and Classical Sanskrit. The term “Sanskrit” is sometimes used as a cover term encompassing the idioms of both the Vedic (= Vedic Sanskrit) and post-Vedic periods.

The absolute chronology of these periods poses serious problems (see, e.g., Witzel 1995: 97f.), so that we can only afford very rough approximations. Thus, the early Vedic period cannot be dated earlier than 1500 BC (and it hardly begins much later than 1200 BC); the middle Vedic period probably starts after 800 BC; and the post-Vedic period must have started somewhere in the second half of the first millennium BC, hardly much earlier than 300 BC.

The Middle Indo-Aryan époque lasts approximately from 600 BC (the time to which the oldest Middle Indic, Pāli, texts may reach back) till the end of the first millennium AD. From 1000 AD onward, the earliest forms of New Indo-Aryan (NIA) languages, such as Old Hindi, Old Bengali etc., are attested.

The general chronology of the Indo-Aryan languages as well as the main subdivisions within the Old and Middle Indo-Aryan periods is shown in Figure 4.1.

Periods Languages	Old Indo-Aryan			Middle Indo-Aryan			New Indo-Aryan	
	P ā l i			Aśoka Prākritis, Middle Prākritis, Apabhraṃśa			Hindi, Marathi, Bengali, Assamese, Sinhala . . .	
	<i>S</i>	<i>a</i>	<i>n</i>	<i>s</i>	<i>k</i>	<i>r</i>	<i>i</i>	<i>t</i>
	Early Vedic	Middle Vedic	Late Vedic	Early Post Vedic – Epic Sanskrit – (Class. Sanskrit)				
<b>Texts</b>	Atharvaveda	Yajurveda, early Brāhṃaṇas	late Brāhṃaṇas, Āraṇyakas	Mahābhārata, Rāmāyana Upaniṣads, Sūtras				
	Ṛgveda							
<b>Years</b>	-1200	-1000	-500	-300	0	500	1000	2000

FIGURE 4.1 CHRONOLOGY OF INDO-ARYAN LANGUAGES AND TEXTS

The sociolinguistic situation in ancient India is a particularly complex issue and requires at least a short commentary. Already by the Middle Vedic period, Sanskrit was no longer a spoken language but co-existed, as a sacral language, alongside the Middle Indo-Aryan (MIA) vernaculars. During the middle and late MIA period, a number of languages (or, to be more precise, “forms of speech”) were used in India. In fact, we are dealing with triglossia, or even polyglossia: Sanskrit was used in the Hindu sacral context, in scientific treatises and some literary works; MIA languages (Prākritis) were used in poetry and dramatic works as well as in religious (Buddhist and Jainist) texts and in epigraphy. Late MIA vernaculars (Apabhraṃśa Prākritis) found their place in the literary tradition as well, while, finally, the colloquial vernaculars, which represented the earliest forms of the NIA languages, were employed in everyday life.

It is important to emphasize that in the course of these developments Sanskrit and Prākritis were not replaced and ousted by later varieties (i.e. Sanskrit by Prākritis, Prākritis by Apabhraṃśas, etc.), but moved up vertically into the position of the high/prestigious form of speech (as indicated by simple arrows in Figure 4.2),<sup>1</sup> to be imitated by the low varieties of speech.

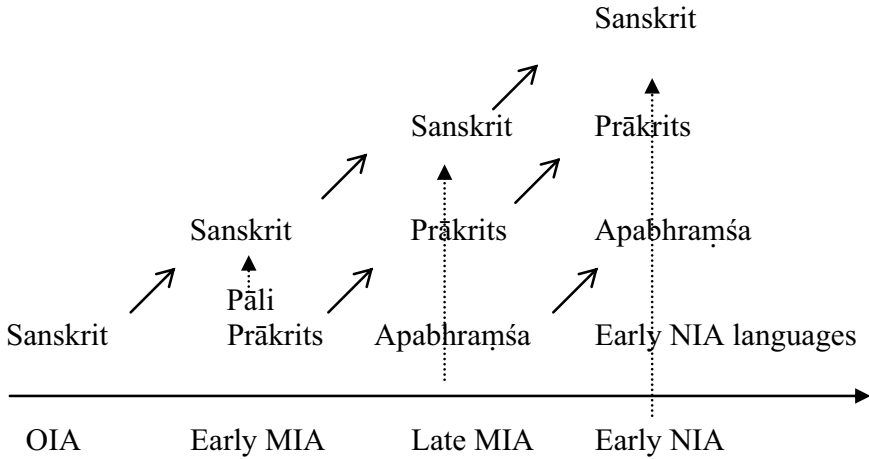


FIGURE 4.2 POLYGLOSSIA IN ANCIENT INDIA (adapted from Bubenik 1998)

All these languages (or forms of speech) co-existed with each other. Most importantly, while the phonological systems and inventories of morphological forms of OIA and MIA languages have been preserved basically intact over the centuries, we can observe numerous traces of the influence of the spoken MIA and NIA vernaculars in the **syntax** and **semantics** of forms in the languages of higher rank. In a way, their grammatical systems, albeit morphologically stable, were open for syntactic “infection” from below, as indicated by the vertical dotted arrows in Figure 4.2. This fact is of crucial importance for understanding the syntactic developments in the late OIA and MIA texts.

Another peculiar feature of the sociolinguistic situation in Ancient India is the enormous authority of the Ancient Indian linguistic tradition, differing in many respects from the younger traditions of Europe and, particularly, that of Pāṇini, the author of the famous grammatical treatise *Aṣṭādhyāyī* (lit. ‘consisting of eight chapters’), dating to the fifth to sixth century BC. By now, it has become the *communis opinio* that the language described by Pāṇini (Pāṇini’s object language) can be roughly identified with middle Vedic Sanskrit (also known as the language of the Vedic prose), attested in the Brāhmaṇas, Āraṇyakas, early Upaniṣads and Sūtras. The oldest of these prose texts can probably be dated to the middle of the first millennium BC (see Figure 4.1). However, this scheme is imprecise in some respects. In particular, we find linguistic phenomena (forms, constructions etc.) that are prescribed by Pāṇini’s grammar but are *not* found in the Vedic corpus. The most plausible explanation for this mismatch can be found in the peculiar sociolinguistic situation in Ancient India, briefly outlined above. Specifically, a number of linguistic phenomena described by grammarians did not appear in Vedic texts but existed within the semi-colloquial scholarly discourse of the learned community of Sanskrit scholars (comparable to Latin scholarly discourse in medieval Europe); for a discussion of the linguistic status of the Pāṇinian Sanskrit, see Kulikov 2013 and Hock 2016. Some such phenomena may result from the influence of Middle Indic dialects spoken by Ancient Indian scholars, thus representing syntactic and morphological calques from their native dialects onto the Sanskrit grammatical system.

Furthermore, we even have some reasons to assume that the rise and rapid development of the Pāṇinian prescriptive grammatical tradition was due, foremost, to the fact that Vedic had ceased to be a living language and the necessity of its codification had been clearly formulated by the contemporary scholarly community. This task was particularly pressing in view of increasing variation within the (semi-colloquial) idiom – essentially based on Middle Vedic Sanskrit, but heavily influenced by Middle Indic dialects – that was used by Ancient Indian paṇḍitas in their scientific and, to some extent, informal discourse.

### Remarks on the genetic classification of (new) Indo-Aryan languages

The genetic classification of Indo-Aryan languages is a much debated and fairly controversial issue. Due to a variety of circumstances, such as the peculiar sociolinguistic situation in Ancient and Medieval India, strong influence from the languages of the sacral and cultural tradition (above all, Sanskrit), and strong convergent processes that never stopped during the more than three thousand year history of the Indo-Aryan languages on the Indian subcontinent, a consistent genetic classification of the modern Indo-Aryan languages is virtually impossible. All classifications are mainly based on structural (grammatical) features, rather than on phonetic developments, and typically posit four or five major subgroups. Thus, according to the notorious Grierson's division, the eastern languages (foremost, Bengali, Oriya and Assamese) share more features with the southern (Marathi and Sinhala with the closely related Maldivian) and northwestern (Sindhi, Lahnda) languages than with two other groups, the "inner", or central (Western Hindi), and the "intermediate" (Eastern Hindi, Panjabi, Nepali etc.), which implies a larger subdivision into "outer" (= eastern + southern + northwestern) and "inner" languages. Some scholars attempted to trace at least this macro-division back to the dialectal variation within late Old Indo-Aryan and MIA periods and consider such features as, in particular, the merger of the three Old Indo-Aryan sibilants (*s*, *ś*, *ṣ*) in *s* as "East-West" (= "outer") innovation. The three- or two-fold macro-division of the Indo-Aryan languages is often connected with the hypothesis about two waves of Indo-Aryan migration, where the first wave is said to correspond to the outer group.

This chapter will concentrate on the earliest attested Indo-Aryan language, Vedic Sanskrit, which can be roughly identified with Old Indo-Aryan and provides the most valuable evidence for Indo-European comparative studies and Indo-European reconstruction. Vedic is one of the most ancient attested Indo-European languages, and from the structural point of view this language is considered in many respects as one of the closest approximations to the reconstructed Proto-Indo-European language. Vedic and post-Vedic Sanskrit are documented with an enormous corpus of literature – larger than the corpus of any other Indo-European literary tradition.

The Vedic corpus includes four main categories of texts, or Vedas, which fulfill different roles in the Hindu ritual: reciting hymns (ṛc-), pronouncing sacrificial formulae (yajus-) that accompany performing sacrifices, chanting songs (sāman-); the fourth Veda, which mainly contains magic spells and is called Atharvaveda was added later to the Vedic canon. Each Veda is transmitted in a number of schools, or "branches" (śākhās), and is subdivided into four genres of texts, depending on their function within the ritual: Samhitās (mantras addressed directly to the gods), Brāhmaṇas and Āraṇyaka (explanatory texts and commentaries on mantras) and Upaniṣads (philosophical texts). These four categories are followed by a variety of Sūtras related to the rituals as well as scientific and law texts (śāstras). The content of the Vedic and early post-Vedic corpus of texts is summarized in Table 4.2.

TABLE 4.2 VEDIC TEXTS AND SCHOOLS

	ṛc- (hymns)		sāman- (chants)		yajus-		
	Śākala	Bāṣkala	Kauthuma	Jaiminiya	White YV		
					Mādhyandina	Kāṇva	
śruti	Samhitās (mantras)	RV	SV	SVJ	VS(M)	VSK	
	Brāhmaṇas	AB	KB	PB ṢB	ŚB(M)	ŚBK	
	Āraṇyakas	AĀ	ŚA		JUB	ŚBM 14.1–8 ŚBK 17	
	Upaniṣads	AiU	KauṣU (= AĀ 2.4–6)	ChU	KenaU (=JUB 4.18–21)	BAUM (= ŚB 14) ĪsU(M) (= VS 40)	BAU(K)
	Śrauta-Sūtras	ĀśŚS	ŚŚS	LatyŚS DrāhyŚS	JŚS	KātyŚS	
vedāṅga	Gṛhya-Sūtras	ĀśGS	ŚGS KauṣGS	GobhGS DrāhyGS	JGS	PārGS	
	Pitṛmedha- Sūtras			ManB, SVB	GautPS		
	Vyākaraṇa [grammatical treatises]: Pāṇini's Aṣṭādhyāyī; Nirukta [etymology]; Jyotiṣa [astronomy/ astrology]; etc.						
Dharma-Sūtras							
smṛti (Dharma-Śāstras)			GautDhS	YājñSmṛ.			

(sacrificial formulae)				atharvan- (magic spells)		
Black YV				Śaunakīya	Paippalāda	
Taittirīya	Maitrāyaṇī	Kāthaka	Kapiṣṭhala			
TS <sup>(m)</sup>	MS <sup>(m)</sup>	KS <sup>(m)</sup>	KpS <sup>(m)</sup>	AV(Ś)	AVP	early Vedic
(TS <sup>p</sup> ), TB	(MS <sup>p</sup> )	(KS <sup>p</sup> ), *Kāth.-Br.	(KpS <sup>p</sup> )		GB	
TĀ		KāthĀ				middle Vedic
TaiU (= TĀ 7–9)						
MNU (= TĀ 10)	MaitrU	KāthU			MuṇḍU	
ŚvetU						
BaudhŚS						
VādhS						
BhārŚS	MānŚS					
ĀpŚS	VārŚS				VaitS	late Vedic
HirŚS						
VaikhŚS						
BaudhGS						
BhārGS						
ĀpGS	MānGS					
HirGS	VārGS	KāthGS			KauśS	
VaikhGS						
ĀgnivGS						
ĀpM						
BaudhPS						
HirPS						
ĀpDhS	ManuSmṛ					
BaudhDhS						
HirDhS						
ViṣṇuSmṛ.						



TABLE 4.3 ABBREVIATIONS OF TEXTS (TEXT SIGLA)

AĀ	Aitareya-Āraṇyaka
AB	Aitareya-Brāhmaṇa
ĀgnivGS	Āgniveśya-Gṛhya-Sūtra
AitU	Aitareya-Upaniṣad
ĀpDhS	Āpastamba-Dharma-Sūtra
ĀpGS	Āpastamba-Gṛhya-Sūtra
ĀpM	Āpastamba-Mantrapāṭha
ĀpŚS	Āpastamba-Śrauta-Sūtra
ĀśGS	Āśvalāyana-Gṛhya-Sūtra
ĀśŚS	Āśvalāyana-Śrauta-Sūtra
AV	Atharvaveda
AVP	AV, Paippalāda recension
AVŚ	AV, Śaunakīya recension
BaudhDhS	Baudhāyana-Dharma-Sūtra
BaudhGS	Baudhāyana-Gṛhya-Sūtra
BaudhPS	Baudhāyana-Pitṛmedha-Sūtra
BaudhŚS	Baudhāyana-Śrauta-Sūtra
BĀU(K)	Bṛhad-Āraṇyaka-Upaniṣad (Kāṇva recension)
BĀUM	Bṛhad-Āraṇyaka-Upaniṣad, Mādhyandina recension
BhārGS	Bhāradvāja-Gṛhya-Sūtra
BhārPS	Bhāradvāja-Pitṛmedha-Sūtra
BhārŚS	Bhāradvāja-Śrauta-Sūtra
Br.	Brāhmaṇas
ChU	Chāndogya-Upaniṣad
DrāhyGS	Drāhyāyaṇa-Gṛhya-Sūtra
DrāhyŚS	Drāhyāyaṇa-Śrauta-Sūtra
GautDhS	Gautama-Dharma-Sūtra
GB	Gopatha-Brāhmaṇa
GobhGS	Gobhila-Gṛhya-Sūtra
HirDhS	Hiraṇyakeśi-Dharma-Sūtra
HirGS	Hiraṇyakeśi-Gṛhya-Sūtra
HirPS	Hiraṇyakeśi-Pitṛmedha-Sūtra
HirŚS	Hiraṇyakeśi-Śrauta-Sūtra
ĪśU	Īśa-Upaniṣad
JB	Jaiminīya-Brāhmaṇa
JGS	Jaiminīya-Gṛhya-Sūtra
JŚS	Jaiminīya-Śrauta-Sūtra
JUB	Jaiminīya-Upaniṣad-Brāhm.
KāṭhĀ	Kāṭha-Āraṇyaka
KāṭhGS	Kāṭhaka-Gṛhya-Sūtra
Kāṭh-Saṃk.	Kāṭhaka-Saṃkalana
KāṭhU	Kāṭha-Upaniṣad
KātyŚS	Kātyāyana-Śrauta-Sūtra
KauśGS	Kauṣītaka-Gṛhya-Sūtra
KauśU	Kauṣītaki-Upaniṣad
KauśS	Kauśika-Sūtra
KB	Kauṣītaki-Brāhmaṇa (= Śāṅkhāyana-Brāhmaṇa)
KenaU	Kena-Upaniṣad
KpS	Kapiṣṭhala-Kāṭha-Saṃhitā
KS	Kāṭhaka(-Saṃhitā)
LātyŚS	Lātyāyana-Śrauta-Sūtra
LaugGS	Laugākṣi-Gṛhya-Sūtra
MaitrU	Maitri- (Maitrī-), Maitrāyaṇa-, Maitrāyaṇīya-Upaniṣad
ManB	Mantra-Brāhmaṇa
MānGS	Mānava-Gṛhya-Sūtra
MānŚS	Mānava-Śrauta-Sūtra
ManuSmṛ.	Manu-Smṛti (= Mānava-Dharma-Śāstra)

---

MNU	Mahā-Nārāyaṇa-Upaniṣad
MS	Maitrāyaṇī Samhitā
MuṇḍU	Muṇḍaka-Upaniṣad
NārSmṛ.	Nārada-Smṛti
Pāṇ.	Pāṇini (Aṣṭādhyāyī)
PārGS	Pāraskara-Gṛhya-Sūtra
PB	Pañcaviṃśa-Brāhmaṇa (= Tāṇḍyamahā-Brāhmaṇa)
PraśU	Praśna-Upaniṣad
Rām.	Rāmāyaṇa
RV	Ṛgveda
RVKh.	Ṛgveda-Khilāni
Sū.	Sūtra(s)
SUB	Samhitopaniṣad-Brāhmaṇa
SV	Sāmaveda (Kauthuma rec.)
SVB	Sāmavidhāna-Brāhmaṇa
SVJ	Sāmaveda, Jaiminiya rec.
ŚĀ	Śāṅkhāyana-Āraṇyaka
ŚB(M)	Śatapatha-Brāhmaṇa (Mādhyandina recension)
ŚBK	Śatapatha-Brāhmaṇa, Kāṇva recension
ŚGS	Śāṅkhāyana-Gṛhya-Sūtra
ŚŚS	Śāṅkhāyana-Śrauta-Sūtra
ŚvetU	Śvetāśvatara-Upaniṣad
ŚB	Śaḍviṃśa-Brāhmaṇa
TĀ	Taittirīya-Āraṇyaka
TaiU	Taittirīya-Upaniṣad
TB	Taittirīya-Brāhmaṇa
TS	Taittirīya-Samhitā
VādhS	Vādhūla-Sūtra
VaikhDhS	Vaikhānasa-Dharma-Sūtra
VaikhGS	Vaikhānasa-Gṛhya-Sūtra
VaikhŚS	Vaikhānasa-Śrauta-Sūtra
VaitS	Vaitāna-Sūtra
VārGS	Vārāha-Gṛhya-Sūtra
VārŚS	Vārāha-Śrauta-Sūtra
VāsDhS	Vāsiṣṭha-Dharma-Sūtra
ViṣṇuSmṛ.	Viṣṇu-Smṛti
VSK	Vājasaneyi-Samhitā, Kāṇva recension
VS(M)	Vājasaneyi-Samhitā (Mādhyandina recension)
YājñSmṛ.	Yājñavalkya-Smṛti
YV	Yajurveda(-Samhitā) (= VS(K), MS, KS, KpS, TS)

---

## PHONOLOGY

### The phonological system of Old Indo-Aryan

The phonological system of Vedic includes only two vowels proper, *a* and *ā*, which differed not only in length but also in timbre (the long vowel was a back vowel). Other monophthong items represent either vocalic allophones of sonants (*y/i*, *r/ṛ* etc.; the contrast between *v* and *u* is phonemic, not allophonic, however) or former diphthongs (*e*, *o*, phonetically long and probably still preserving the diphthong pronunciation in early Old Indo-Aryan). The vowel system can be summarized as follows:

Vowels proper:

a ā

Diphthongs:

e o  
ai au

Sonants that have vocalic variants:

v ~ u ū  
y / ī ī  
r / ṛ (ṛ)  
l / (ḷ)

The consonant system includes plosives (voiceless, voiceless aspirates, voiced and voiced aspirates) and nasals (some of which have weak or no phonemic status), organized by five places of articulation, as well as three sibilants and the voiced (pharyngeal) *h*:

**TABLE 4.4 OLD INDO-ARYAN CONSONANTISM**

pharyngeal	(morpho- phonological symbols)	Place of articulation					laryngeal
		labial	dental	retroflex	palatal	velar	
Voiceless stops	(T)	p	t	ʈ	c	k	
Voiceless aspirates	(Th)	ph	th	ʈh	ch	kh	
Voiced stops	(D)	b	d	ɖ/ḍ	j	g	
Voiced aspirates	(Dh)	bh	dh	ɖh	(jh)	gh	
Nasals	(N)	m	n	ɳ	[ɲ]	(ṅ)	
Voiceless fricatives	(S)		s	ʂ	ś		
Voiced fricatives							h

### Accentuation

A number of the early and middle Vedic texts mark accents. In particular, both early Vedic texts, RV and AV (in the Śaunakīya recension), as well as a number of early Saṃhitā prose texts (Taittirīya and Maitrāyaṇī) and early Brāhmaṇas (Taittirīya, ŚB), are preserved in accented form.

Phonetically, three tones (degrees of pitch) can be distinguished, high (udātta´) [= principal accent], middle (svarita´) and low (anudātta). Phonologically, we only have an opposition between the vowels (or syllables) that bear the principal accent, or the high pitch, and those which bear no principal accent (= unaccented), i.e. middle or low pitch. A word-form can have only one principal accent (or high pitch), except for some compounds and *tavái*-infinitives. The placement of middle and low pitches is determined automatically within the sequence of words forming metrical units (pāda): (i) all vowels/syllables that have no principal accent bear the low pitch (anudātta); and (ii) the vowel/syllable that immediately follows udātta receives the middle pitch (svarita), unless the following syllable bears high pitch (udātta).

The European tradition of the transcription of Vedic texts marks only the principal accent (high pitch),<sup>2</sup> using the acute accent mark (*á* etc.). Svarita is marked, by using the gravis accent mark (*à* etc.), only if the vowel that bears the preceding udātta is not written as the full vowel as well as in the case of “independent svarita” (due to a particular

external sandhi). By contrast, the Devanāgarī notation accepted for the RV, AV and Taittirīya texts marks only udātta and svarita.

There are certain historically conditioned (essentially going back to the Proto-Indo-European morphophonological system) correlations between the ablaut grade and the placement of the accent. In particular, the first morpheme that shows a non-zero grade typically bears the accent. Yet numerous exceptions and analogical developments make these two parameters essentially independent from the synchronic point of view.

A number of word-forms bear no principal accent. These include (i) finite verbal forms, unless they are employed at the beginning of a sentence and/or pāda (= minimal metrical unit), or in a subordinate clause; (ii) vocatives, except at the beginning of a sentence and/or pāda; and (iii) clitics.

## History of common Indo-Aryan phonemes

### Vowels and diphthongs

The vowel system of Common Indo-Aryan is almost identical to that of Proto-Indo-Iranian and reproduces its main distinctive features: the loss of most quality distinctions (*e/o*), as opposed to the preservation of most quantity (length) distinctions. The main innovations in Indo-Aryan (partly shared with Iranian) include (i) the merger of all proper Proto-Indo-European vowels (that is, *e* and *o*) as well as the vocalic allophones of the nasal sonants in *a*; the same holds for its long pendants and vocalic elements of diphthongs; (ii) monophthongization of diphthongs (*Vī* > *e*, *Vu* > *o*); and (iii) merger of the vocalic variants of *i/y* and of all laryngeals in *i*.

The history of the Common/Old Indo-Aryan vowels and diphthongs can be illustrated with the following examples:

#### *a* < \**e*, \**o*, \**ṃ*, \**ṅ*

*dāśa* ‘ten’ < \**dekṃ*, cf. Gr. δέκα, Lat. *decem*

*mārta-* ‘mortal, man’ < \**mor-to-*, cf. Gr. (Kallimachos) μορτοί pl.

*ājra-* ‘field’ < \**h<sub>2</sub>eǵ-ro-*, cf. Gr. ἀγρός, Lat. *ager*, Goth. *akrs*

*āsthi-* n. < \**h<sub>3</sub>estH-* ‘bone’, cf. Hitt. *ḥastai-* l, Gr. ὀστέον, YAv. *ast-* n. ‘bone, body’

*matī-* f. ‘thought’ < \**ṃṅ-ti-*, cf. Lat. *mēns, mentis* f. ‘mind’, Lith. *mintis* f.

*a-(pútra-)* ‘without (a son)’ < \**ṅ-* adj., cf. YAv. *a-puθra-*, Gr. ἄ-θεος adj. ‘without a god’, Lat. *in-*, Goth. *un-*

#### *ā* < \**ē*, \**ō*, \**o* (| *\_CV, \_#*), \**eH*, \**oH*, \**ṃH*, \**ṅH*

*mātā* ‘mother’ < \**meh<sub>2</sub>tē(r)*, cf. Av. *mātā*, Gr. μάτηρ, Lat. *māter*

*vāk* ‘speech, voice’ < \**wōk<sup>w</sup>s*, cf. OAv. *vāxš*, Lat. *uōx*

*jānu-* ‘knee’ < \**ǵónu-*, cf. YAv. *zānu* °, Gr. γόνυ

The only trace of the original Proto-Indo-European \**e/o* distinction can be found in the open non-final syllable, where PIE \**o* is lengthened, in accordance with Brugmann’s Law, which operated as early as Indo-Iranian, cf. 1 sg. perf. *jagáma* < \**g<sup>w</sup>e-g<sup>w</sup>om-h<sub>2</sub>e* (cf. Gr. μέμονα) vs. long vowel in 3 sg. perf. *jagāma* < \**g<sup>w</sup>e-g<sup>w</sup>om-e* (cf. Gr. μέμονε); or in causatives like *janáyati* ‘to beget’ (< \**ǵonh<sub>2</sub>-eye-*), not \*\**jānáyati!*)

*jātá-* ‘born’ < \**ǵṅH-to-*, cf. Lat. *nātus*

**i < \*i, \*H/ə (ir < \*rH | \_V)**

*i-hí* ‘to go’, 2 sg. imp. < \**h<sub>1</sub>i-d<sup>h</sup>i*, cf. OCS *idi*, Gk. ἴθι

*pítár-* ‘father’ < \**ph<sub>2</sub>té(r)*, gen. \**ph<sub>2</sub>tr-és*, \*-ós, cf. Gr. πατήρ, πατρός; Lat. *pater*, *-tris*;

Goth. *fadar*

*-mahí* (sec. ending 1 pl. mid.) < \**med<sup>h</sup>h<sub>2</sub>*, cf. Gr. -μεθα

*-tirá* ‘to pass’ 2 sg. imp. act. < \**-tṛh<sub>2</sub>-e-*, cf. LHitt. *tar-ra-at-ta* (< \**térh<sub>2</sub>-o-to-*) ‘to be able’; Lat. *trāns* (< \**tṛh<sub>2</sub>-nt-*)

**ī < \*iH, \*H/ə (īr < \*rH, \*iH | \_C)**

*jīvá-* ‘alive, living’ < \**g<sup>w</sup>i<sub>h</sub>₃-wó-*, cf. Lat. *uīuus*, OCS *živъ*, Lith. *gývas*

*pu-n-ī-mas* ‘to cleanse’ 1 pl. pres. < \**pu-n-H-mes*

*a-tārūt* ‘to pass’ 3 sg. sigm. aor. < \**e-tēr<sub>h</sub>₂-s-*

*tīrṇa-* ‘to pass’ perf. ptcp. < \**-tṛh<sub>2</sub>-no-*

**u < \*u (ur < \*rH | \_V)**

*vásmi*, *uśmási* ‘to wish’, 1 sg., 1 pl. pres. < \**wek-mi*, \**uk-mes*, cf. OAv. *vasamī*,

*usəmahi* /*vasmi*, *usmahi*/, Hitt. 1 sg. *ú-e-ek-mi*

*sphuráti* ‘kicks’ < \**sp<sup>(h)</sup>rH-e-*, cf. Lat. *spernō* ‘to dissociate, reject, spurn’, Oic. *sperna* ‘to kick out with the feet’

**ū < \*uH (ūr < \*rH, \*iH | \_C)**

*múš-* ‘mouse’ < \**muHs-*, cf. Gr. μῦς; Lat. *mūs*; OHG *mūs*; Russ. *мышь*

*pūrṇá-* ‘full’ < \**plH-no-*, cf. Goth. *fulls*, Lith. *pilnas*, OCS *plъnъ*, Av. *pərəna-*

**ʀ < \*r, \*l**

*mṛtá-* ‘died, dead’ < \**mṛ-to-*, cf. Lat. *mortuus*

*vṛka-* ‘wolf’ < \**wṛk<sup>w</sup>o-*, cf. Goth. *wulfs*, Lith. *vilkas*, ORuss. *вѣлкѣ*

*ř* arises from *ʀ* analogically, cf. *piřn* ‘father’, acc. pl. (\**ph<sub>2</sub>tṛ-ns*) in analogy with *devá-* ‘god’: acc. pl. *devān*

**e < \*ey, \*oy**

*éti* ‘to go’, 3 sg. pres. < \**h<sub>1</sub>eyti*, cf. Gr. εἶσι

*cetáyati* ‘to reveal’ < \**k<sup>w</sup>oyt-eye-*, cf. Goth. *haidu-* (< \**k<sup>(w)</sup>oyt-ú-* or \**k<sup>(w)</sup>oy-tú-*) ‘manner’

**ai < \*ēy, \*ōy**

(*vṛk-*)*aiḥ* ending instr. pl. *o*-decl. < \**-ōys*, cf. Lith. *vilk-aĩs*, Lat. *lup-īs*

**o < \*ew, \*ow**

*ójas-* ‘strength, power’ < \**h<sub>2</sub>ewg-es-*, cf. Lat. *augustus* ‘elevated’; Lith. *augestis* ‘growth’

*bódhati* ‘to perceive’ < \*b<sup>h</sup>ewd<sup>h</sup>-e-, cf. Gr. πεύθομαι ‘to learn, to hear’  
*bodháyati* ‘to make awake’ (caus.) < \*b<sup>h</sup>owd<sup>h</sup>-eye-ti, cf. OCS *buditi*

**au** < \*ēw, \*ōw

*dyáuh* ‘heaven, Father Sky’ nom. sg. < \*dyēw-s, cf. Gr. Ζεύς, Lat. *diūs*  
*gáuh* nom. sg. ‘cow’ < \*g<sup>w</sup>ōw-s, cf. Gr. βούς

### Sonants

The system of sonants is quite conservative, except for the weakening of the phonemic status of *l* (completely lost in Iranian, but partly preserved in Indo-Aryan), which merges with *r* in many cases.

**y** < \*i/y | \_V (allophonic variant of *i*)

*yúvan-* ‘young, young hero’ < \*h<sub>2</sub>yu-h<sub>1</sub>on-, cf. Arm. *yavanak/yovanak* ‘young (of animal)’; Lat. *iuvenis*

**r** < \*r, \*l

*bhrú-* ‘(eye)brow’ < \*h<sub>3</sub>b<sup>h</sup>ruH-, cf. Gr. ὄφρῦς; OEng. *brū*; Lith. *bruvìs*; OCS *brъvb*  
*mriyá-te* ‘die’ < \*mṛ<sub>1</sub>-ye-, cf. Lat. *moriór*  
*śrávas-* ‘fame’ < \*kléwes-, cf. Gr. κλέος ‘call, fame’, OCS *slovo* ‘word’

**l** < \*l, (\*r)

*lúbhyati* ‘to be in disorder’ < \*lub<sup>h</sup>-, cf. OCS *ljubiti* ‘to love’  
*reh-/leh-* (RV)/(AVP+) ‘to lick’ < \*leyǵ<sup>h</sup>-, cf. OCS *lizati*

**v** < \*u/w | \_V

*vák* ‘speech, voice’ < \*wōk<sup>w</sup>s, cf. OAv. *vāxš*, Lat. *uōx*

### Plosives

The system of sonants and plosive consonants is both conservative, preserving the main contrasts of the Proto-Indo-European consonantism, and innovative, exhibiting a number of additions to the original system. The most important innovations include (i) the rise of voiceless aspirates from the combination of voiceless stops and laryngeals (\*TH > Th); (ii) the rise of retroflex consonants, foremost on the basis of retroflexivization of *s* after *r*, *k* and vowels distinct from *ā* (see also the section on sandhi below), which further retroflexivizes adjacent dental stops; retroflexivization of *n* after *r*; and the development of final \*-ks > *t*; (iii) two palatalizations, which result, in particular, in the merger of the Proto-Indo-European voiced palatal gutturals (\*ǵ and \*ǵ<sup>h</sup>) with simple voiced gutturals before front vowels (*e*, *i*) in *j* and *h*, respectively; notice that these two types of sources can still be distinguished morphophonologically; that is, they yield distinct reflexes (*j*<sub>2</sub> and *h*<sub>2</sub> as opposed to *j*<sub>1</sub> and *h*<sub>1</sub>) in sandhi.

The history of the Common/Old Indo-Aryan consonants is summarized and illustrated with examples below:

### VELARS

**k** < \***k**<sup>(w)</sup> |  $\_X_{\neq *ē, *i}$

*kád* nom./acc. sg. n. interrog. pron. < \***k**<sup>w</sup>od, cf. Lat. *quod*, OHG *hvaz*

*loká-* m. ‘free/light space, world’ < \***k**<sup>w</sup>ewk-o-, cf. Gr. λευκός ‘light, white, bright’, Lat. *lūx* f. ‘light’

**kh** < \***k**<sup>(w)</sup>H |  $\_V_{\neq *ē, *i}$

*śákhā-* ‘branch, twig’ < \***k**<sup>w</sup>[k<sup>w</sup>]-h<sub>2</sub>-, cf. Goth. *hoha* ‘plough’, Lith. *šakà* ‘twig’, ORuss. *soxa* ‘(wooden) plough’

**g** < \***g**<sup>(w)</sup> |  $\_X_{\neq *ē, *i}$

*gám* acc. sg. ‘cow’ < \***g**<sup>w</sup>ōm, cf. OAv. *gəm*, Gr. βῶν, Umbr. *bum*

**gh** < \***g**<sup>(wh)</sup> |  $\_X_{\neq *ē, *i}$

*ghnánti* ‘to slay’ 3 pl. pres. < \***g**<sup>wh</sup>nenti, cf. Hitt. *ku-na-an-zi*

**ñ** < \***n** |  $\_C_{\text{velar}}$

The velar nasal emerges only through sandhi before velar plosives, and the following example represents the only grammatical context where this sound has phonemic status: *prāñ* nom. sg. m. of the adjective *prāñc-* ‘directed forwards’ ← *prāñk-s* ← *prāñc-s* (sandhi)

### PALATALS

**c** < \***k**<sup>(w)</sup> |  $\_ *ē, *i$

*ca* ‘and’ < \***k**<sup>w</sup>e, cf. Gr. τε, Lat. *que*

*cakrá-* ‘wheel’ < \***k**<sup>w</sup>e-k<sup>w</sup>l-o-, cf. Gr. κύκλος ‘circle, ring, wheel’; Toch. A *kukäl* ‘cart’

**ch** < \***sk** |  $\_ *ē, *i/y$

*chid-* ‘to split, break’ < \***sk**id-, cf. YAv. *sið-*, Gr. σχίζω ‘to split, cut’, Lat. *scindō* ‘to cut open’

*gáčhati* pres. stem ‘to go’ < \***g**<sup>w</sup>ṃ-ske-, cf. YAv. *jasaiti* 3 sg. pres., Gr. βάσκε 2 sg. imp. act. ‘go!’

**j**<sub>(1)</sub> < \***g**<sup>(w)</sup> |  $\_ *ē, *i$  ( $j_1 = \{j/g/k\}$ )

*jīvá-* ‘alive, living’ < \***g**<sup>w</sup>i<sub>h</sub>-wó-, cf. Lat. *uīuus*, OCS *živъ*, Lith. *gývas*

*yuj-* ‘to connect, yoke’, *yúj-* ‘connected, yoked; companion’ < \***Hyug-** ‘to yoke, join’ (e.g. in dat.sg./inf. *yuj-é* < \***Hyug-ei**, ins.sg. *yuj-ā* < \***Hyug-eh**<sub>1</sub>) (: *yuk-tá-* perf. ptcp. < \***Hyug-to-**), cf. Gr. ζεύγνυμι, Lat. *iungere*, Lith. *jūngti*, OCS *igo* ‘yoke’

**j**<sub>(2)</sub> < \***ǵ** ( $j_2 = \{j/š\}$ )

*jānu-* ‘knee’ < \***ǵ**ónu-, cf. YAv. *zānu* °, Gr. γόνυ

*yaj-* ‘to worship, sacrifice’ < \***Hyaǵ-** (cf. ptcp. perf. *iš-tá-* < \***Hiǵ-to-**), cf. Gr. ἄζομαι ‘to honour’, Gr. ἅγιος ‘sacred, consecrated’

*jh* emerges only through borrowing or in onomatopoeia, as in *jájjhat-* pres. ptcp. ‘laughing’

[*ñ*] < *n* |  $\_C$ <sub>palatal</sub>, *ñ* is a non-phonemic (allophonic) variant of *n*, cf. *yu-ñ-j-anti* ‘to yoke, join’ 3 pl. pres. < \*Hyu-n-g-nti

## RETROFLEX

*ṣ* < \**s* | \**i*, \**u*, \**r*, \**k*  $\_$

*ṣ* < \**K*, \**ḡ* |  $\_$ \**t*

*áraiṣam* ‘to leave’ 1sg. aor. < \*(-leyk<sup>w</sup>-)s-, cf. Gr. ἔλευσα

*viśá-* ‘venom, poison’ < \*wis-, cf. Gr. ἰός, Lat. *uīrus*

*juṣ-* ‘to enjoy’ < \*ḡus-, cf. Lat. *gustō* ‘to taste’, Goth. *ga-kiusan* ‘to test’

*várṣman-* ‘height, peak, top’ < \*wers-men-, cf. Lith. *viršùs*, OCS *vrъxъ*

*váṣ-ṭi* ‘to wish, desire’ 3 sg. pres. < \*weḱ-ti, cf. Gr. ἐκόν ‘voluntary’, Hitt. /uēḱtsi/

*ṭ* < *t* |  $\_s$   $\_$ , \**Ks* |  $\_$ #

*juṣṭá-* < \*ḡus-to- ‘enjoy’ perf. ptcp.

*viṭ* (*viś-*) ‘settlement, community’, nom.sg. < \*wik-s, cf. OCS *vbъb* ‘village’, Gr. οἶκος ‘house, dwelling’

*ṭh* < \**tH* |  $\_s$   $\_V$

*tíṣṭha-* ‘stand’, pres. < \*stí-sth<sub>2</sub>-e-, cf. Gr. ἵστημι ‘I place’

*ḍ* < \**ḷ* < \**ṣ* |  $\_D$ <sup>(h)</sup>

*viṣṭá-bhiḥ* < \*-prus-b<sup>h</sup>-. . . ‘drop’ (*vi-pruṣ-*)

[*ḷ*] is an allophonic variant of *ḍ* that emerges only in the intervocalic position (|  $\_V$   $\_V$ ) in the dialect of the RV, cf. RV *īḷe* = (*īḍe* in other texts) ‘I worship’ 1 sg. pres.

( $\_V$ )*ḍh* < \*( $\_V$ )*ḷdh* < \*( $\_V$ )*ṣdh*

*ūḍhá-* ‘carry’ perf. ptcp. ( $\sqrt{vah}$ ) \*uḡ<sup>h</sup>-to- (> \*uḡ<sup>h</sup>-d<sup>h</sup>o-)

*nīḍá-* < \*ni-sdo- ‘nest’, cf. Lat. *nidus*, Eng. *nest*

*ṇ* < \**n* |  $\_s$ , *r* ( . . )  $\_$

Notice that the retroflexivization of *n* is blocked by dental, retroflex and palatal consonants in ( . . ),

cf. *īrṇa-* ‘to pass’ perf. ptcp. < \*ṭr<sub>h</sub>₂-no-

*várṣmāṇ-am* ‘height, peak, top’ < \*wers-men-

*gr̥háyāmāna-* (\*gr̥b<sup>h</sup>-ye-mh₁no-) ‘to seize’ pres. pass. ptcp. but not *rájānām* ‘king’ acc. sg.; *pr̥ṣṭhēna* ‘back’ instr. sg.; etc.

## DENTALS

*t* < \**t*

*tri-* ‘three’ < \*tri-, cf. Gr. τρεῖς, Lat. *trēs* (nom. pl.), Goth. *þrins* (acc. pl. m./f.), OCS *tri* (f.)

*th* < \**tH* |  $\_V$

*pr̥thú-* ‘broad, wide’ < \*pṭh<sub>2</sub>-ú-, cf. Gr. πλατύς, Lith. *platus*



**d < \*d**

*pád-* ‘foot’ < \*ped-/\*pod-, cf. Gr. (Dor.) πῶς, ποδός (gen. sg.), OEng. *fēt* (nom. pl.),  
Lat. *ped-is* (gen. sg.)

**dh < \*d<sup>h</sup>**

*dádhā-ti* ‘to put, place’ 3 sg. pres. < \*d<sup>h</sup>e-d<sup>h</sup>eh<sub>1</sub>-ti, cf. Gr. τίθη(-μι)

**n < \*n**

*náva-* ‘new, young’ < \*newo-, cf. Hitt. *neṃa-*; Gr. νέος; Gr. (Myc.) *ne-wo*; Lat. *nouus*;  
OCS *novъ*

**LABIALS****p < \*p**

*pitár-* ‘father’ < \*ph<sub>2</sub>tér-, nom.sg. *pitá* < \*ph<sub>2</sub>té(r), cf. Gr. πατήρ, Lat. *pater*, Goth. *fadar*

**ph < \*pH | \_V**

*phéna-* ‘foam’ < \*pHoyno- (?), cf. OCS *pěna*

**b < \*b**

*bála-* ‘power, strength’ < \*b<sup>h</sup>el-o-, cf. Lat. *dē-bilis* ‘without strength’, Gr. βέλτερος  
‘better’; OCS *bolijъ* ‘bigger’

**bh < \*b<sup>h</sup>**

*bhuj-* ‘to enjoy, consume’ < \*b<sup>h</sup>ewg-, cf. OAv. *būj-* f. ‘expiation’, Lat. *fungor* ‘to enjoy,  
suffer, get rid of’

*bhrú-* ‘(eye)brow’ < \*h<sub>3</sub>b<sup>h</sup>ruH-, cf. Gr. ὄφρῦς; OEng. *brū*; Lith. *bruvìs*; OCS *brъvъ*

**m < \*m**

*mātā* ‘mother’ < \*meh<sub>2</sub>tē(r), cf. Av. *mātā*, Gr. (Dor.) μᾶτηρ, Lat. *māter*

**SIBILANTS and h****s < \*s**

*as, ásti* 3 sg. act. ‘to be’ < \*h<sub>1</sub>es-, cf. Gr. ἔσσι, ἐστί; Hitt. *e-eš-zi*; Lat. *est*; Goth. *ist*

**ś < \*k**

*dáśa* ‘ten’ < \*dekm̥, cf. Gr. δέκα, Lat. *decem*

*śván-* ‘dog’ < \*kwon-/\*kun-, cf. Arm. *šown*, Gr. κύων, κυνός, OIr. *con* (gen. sg.), Lith.  
*šuo*, Hitt. *kuṃan/kun-* ‘dog-man’, Oİc. *hun-d-r*

**h<sub>(1)</sub> < \*g<sup>(w)h</sup> | \_ \*č̣, \*i**

*hánti* 3 pl. pres. ‘to slay’ < \*g<sup>wh</sup>en-ti

*dáhati* ‘to burn’ 3 sg. pres. < \*d<sup>h</sup>eg<sup>wh</sup>-e-ti (: *dagdhá-* ptcp. perf. < \*d<sup>h</sup>eg<sup>wh</sup>-to-), cf. Lith.  
*degù*, Lat. *foueō* ‘to make warm’

**h<sub>(2)</sub> < \*ǵ<sup>h</sup> (also < d<sup>h</sup>, mostly in | V<sub>#f</sub>-Ṽ; rarely < b<sup>h</sup>)**

*hásta* ‘hand’ < \*ǵ<sup>h</sup>es-to-, cf. Av. *zasta-*; Lith. *pa-žastis* ‘armpit’

*ahám* ‘I’ < \*h<sub>1</sub>eǵH-om, cf. Av. *azəm*, OPers. *adam*, OCS *azъ*, cf. Gr. ἐγώ, Lat. *egō* <  
\*h<sub>1</sub>eǵ-oH

*váhati* ‘to move, carry, drive’ 3 sg. pres. < \*weǵ<sup>h</sup>-e-ti (: *ūdhá-* ptcp. perf. < \*uǵ<sup>h</sup>-d<sup>h</sup>o- < \*uǵ<sup>h</sup>-d<sup>h</sup>o- < \*uǵ<sup>h</sup>-to-), cf. Av. *vazaiti*, Lat. *uehō*, -ere, OIc. *vega*, OHG *wegan*, OCS *vezq ihá* ‘here’ < \*id<sup>h</sup>a, cf. Pāli *idha*, LAv. *iḍa*; cf. also *ihī* ‘to go’, 2 sg. imp. discussed on p. 224.

Note that the voiced sibilants *z* (existing in Avestan) and \**z* disappeared in Proto-Indo-Aryan.

To conclude this section, one should mention two important phonetic laws that are important for the development of aspirates in Old Indian:

Bartholomae’s Law (D<sup>h</sup>+T > D<sup>h</sup>D<sup>h</sup>, D<sup>h</sup>+s > D<sup>h</sup>z<sup>h</sup>, also where the consonants are separated by *s* or a laryngeal) is responsible for the shift of aspiration to the second member of a consonant cluster and its subsequent voicing, cf. *buddhá-* ‘awaken’ < \*b<sup>h</sup>ud<sup>h</sup>-to-; *ápi gdha* 3 sg. inj. mid. ‘to devour’ < PIIr. \*-g<sup>h</sup>ǵd<sup>h</sup>a < PIE \*g<sup>h</sup>s-to; *duhitár-* ‘daughter’ < \*d<sup>h</sup>uǵ<sup>h</sup>Htar- < PIE \*d<sup>h</sup>uǵ<sup>h</sup>ter- (OAv. *dugədar-*).

Grassmann’s Law (D<sup>h</sup> . . . D<sup>h</sup> > D . . . D<sup>h</sup> | . . . ≤ 1 syllable) accounts for the loss of the aspiration of the first aspirate in case of a sequence of aspirates separated by no more than one syllable, cf. *dugh-* ‘to give milk’ < PIE \*d<sup>h</sup>ewǵ<sup>h</sup>- (cf. Goth. *daug* ‘to be good for smth., fit’); *dá-dhā-ti* ‘to put, place’ 3 sg. pres. < PIE \*d<sup>h</sup>e-d<sup>h</sup>eh<sub>1</sub>-ti (cf. Gr. τίθη(-μυ)); *bódh-a-ti* ‘to perceive’ 3 sg. pres. < PIE \*b<sup>h</sup>ewd<sup>h</sup>-e-ti (cf. Gr. πρῆθομαι ‘to learn, to hear’).

## MORPHOPHONOLOGY

### Morphotactics

Sanskrit is notorious for its extremely rich system of processes that occur at morpheme or word boundaries that are called, in accordance with Indian tradition, sandhi (a Sanskrit term, literally meaning ‘together-posing, joining’, which has been borrowed into European linguistic terminology). Some of these processes have an allophonic nature (such as, for instance, palatalization of the dental nasal before palatals) and thus can be described in terms of phonological rules, but many others represent phonological, non-automatic phenomena and thus belong to the domain of morphophonology. The main two types of sandhi include internal and external sandhi. External sandhis operate on the boundaries between words (= free morphemes), as well as, normally, (i) between members of compounds, (ii) between preverbs/prefixes and roots, and (iii) between nominal stems and the following endings of the dual and plural: three plural/dual case endings beginning with *bh-*, i.e. *-bhis*, *-bhyas* and *-bhyām* and the loc. pl. ending *-su*. Internal sandhis operate on the boundaries between (bound) morphemes.

Typical examples of internal sandhi include, for instance:

- i) C<sub>1</sub>C<sub>2</sub> . . . C<sub>n</sub> → C<sub>1</sub> | # (all members of the final consonant cluster are dropped except for the first one), cf. *pácant-s* → *pácan* (act. ptcp. (nom. sg. m.) of *pac* ‘cook’); *á-chānd-s-t* → *áchān* (3 sg. aor. act. of *chand* ‘appear’);
- ii) devoicing and deaspiration of stops before *s* or in auslaut: T(h), D(h) → T | #, \_s [except for palatals: T(h), D(h) ≠ c, j], cf. *labh-syate* → *lapsyate* (3 sg. fut. mid. of *labh* ‘obtain’); *yúdh-s* → *yút* ‘fighter’ (nom. sg.); *yúdh* → *yút* ‘O fighter!’ (voc. sg.);
- iii) retroflexivization of *s* after *r*, *k* and vowels distinct from *ā* (see 2.3.2 above), also known as the “RUKI rule”: *s* → ṣ | k, r, V<sub>#ā</sub> \_ (not \_#, r = not in auslaut or before r), cf. *á-raik-s-am* → *áraiṣam*, *á-kār-s-am* → *ákārṣam*, *juhó-si* → *juhó-ṣi*;

- iv) assimilation to the following plosive in the way of articulation (i.e. T(h), D + t → Tt; T(h), D + th → Tth; T(h), D + dh → Ddh), accompanied by progressive retroflexivization (retroflex consonants retroflexivize the following stop), which can be conveniently presented in table form (Table 4.5; adopted from Zalznjak 1978/1987):

TABLE 4.5 ASSIMILATION IN CONSONANT CLUSTERS  
BEFORE DENTAL STOPS

	+ t	+ th	+ dh
<i>k, kh, g, c, j<sub>1</sub> +</i>	<i>kt</i>	<i>kth</i>	<i>gdh</i>
<i>t, th, d +</i>	<i>tt</i>	<i>tth</i>	<i>ddh</i>
<i>t, th, d +</i>	<i>tt</i>	<i>tth</i>	<i>ddh</i>
<i>p, ph, b +</i>	<i>pt</i>	<i>pth</i>	<i>b dh</i>
<i>ś, ṣ, ch, j<sub>2</sub> +</i>	<i>st</i>	<i>sth</i>	<i>ddh</i>
<i>s +</i>	<i>st</i>	<i>sth</i>	<i>dh</i>

Cf. *yoj<sub>1</sub>-tum* → *yoktum* (inf. of *yuj* ‘yoke, join’); *vac-tum* → *vaktum* (inf. of *vac* ‘say’); *prach-tum* → *praštum* (inf. of *prach* ‘ask’); *mṛj<sub>2</sub>-tha* → *mṛṣṭha* (2 pl. pres. act. of *mṛj* ‘adorn, anoint’); *mṛj<sub>2</sub>-dhvam* → *mṛḍḍhvam* (2 pl. pres. act. of *mṛj* ‘adorn, anoint’); *ās-dhve* → *ādhve* (2 pl. pres. mid. of *ās* ‘to sit’); but cf. also *as-dhi* → *edhi* (2 pl. imp. of *as* ‘to be’).

External sandhi (many of which – especially the vowel sandhi – are optional in early Vedic) can be illustrated, for instance, by the following rules:

- i) merger (contraction) of two homorganic vowels (optional in early Vedic, which is shown by a bracketed arrow): -V + V- → -V̄-, cf. *ihá asti* → *ihāsti* ‘here is’; *su úktám* → *sūktám* ‘well-said’;
- ii) monophthongization:

-ā + ī- → -e- (cf. *ihá iha* → *ihéha*, *pitāiva* → *pitéva* ‘like father’)

and

-ā + ū- → -o- (cf. *á ubhá* → *óbhá* ‘to both’);

-as + a- → -o-, cf. *devás atha* → *devó tha* ‘and the god’.

### Ablaut alternation

Most morphemes may appear in one of the three alternation grades (traditionally called “ablaut” grades): (i) zero, or weak; (ii) full, or normal (*guṇa* “quality” in Indian tradition); and (iii) long (*vṛddhi* ‘increasing’). Historically, this alternation goes back to the Proto-Indo-European ablaut, \*zero/\*e/\*o/\*ē/\*ō. Thus, for roots with the structure *CaC*, the alternation between the zero, full and long grades will manifest as *ø/a/ā* (e.g. *pt-/pat-/pāt-* ‘fly’). The alternating vowel can be followed by a sonant (*i/y, u/v* etc.), which can be realized in different ways, depending on the phonological context. Altogether, this results in a rich allomorphy, such as *i – e/ay – ai/āy* (e.g. *jī-/je-/jay-/jai-/jāy-* ‘win’) etc.

The choice of the alternation grade is determined by the grammatical form. Next to the forms that require one of the three “traditional” grades listed above, there are a number of grammatical positions (causative stem, 3 sg. perf. act.), where some morphemes

show the full grade, while others have the long grade. In Indo-Iranian, *e* and *o* merge in *a*, but *o* changed to *ā* in open syllables, according to Brugmann’s Law. Thus, the original ablaut is reflected in Vedic as the zero/*a/ā* + *a/ā* opposition. Accordingly, for descriptive purposes it seems appropriate to posit (iv) the fourth (“Brugmann’s”) grade for such grammatical contexts as causative stem or 3 sg. perf. act.

Table 4.6 presents the main alternation grades (in a somewhat simplified form; for detailed rules that describe the choice of the alternation grade depending on the phonological context, see Zaliznjak 1975):

**TABLE 4.6 THE MAIN ALTERNATION GRADES**

	a(C)	a(CC)	i(V)	i(C)	u(V)	u(C)	...	ṛ	ṛ(C)	...	an	an(C)	am(V)	...
zero	–	–	i(y)	i	u(v)	u	...	ṛ	ṛ	...	a/n	a/n	a(m)/m	...
(guṇa) full	<i>a</i>	<i>A</i>	<i>ay</i>	<i>e</i>	<i>av</i>	<i>o</i>	...	<i>ar</i>	<i>ar</i>	...	<i>an</i>	<i>an</i>	<i>am</i>	...
(Brugmann’s grade)	<i>ā</i>	<i>A</i>	<i>āy</i>	<i>e</i>	<i>āv</i>	<i>o</i>	...	<i>ār</i>	<i>ar</i>	...	<i>ān</i>	<i>an</i>	<i>ām</i>	...
(vṛddhi) long	<i>ā</i>	<i>ā</i>	<i>āy</i>	<i>ai</i>	<i>āv</i>	<i>au</i>	...	<i>ār</i>	<i>ār</i>	...	<i>ān</i>	<i>ān</i>	<i>ām</i>	...

	a(C)	...	i(C)	u	...	am
zero	<i>á-pa-pt-at</i>	...	<i>rik-tá-</i>	<i>śru-tá-, śrúv-at</i>	...	<i>ya-tá-, yam-et</i>
full	<i>pát-ati</i>	...	<i>á-ri-rec-īt</i>	<i>śro-śyáti</i>	...	<i>yám-itavái</i>
(Brugmann’s)	<i>pāt-áyati</i>	...	<i>ri-réc-a</i>	<i>śrāv-áyati</i>	...	<i>yām-áyati</i>
long	–	...	<i>á-raik</i>	<i>á-śrau-ṣīt</i>	...	<i>á-yām-s-am</i>
	<i>pat</i> ‘fly’		<i>ric</i> ‘leave’	<i>śru</i> ‘hear’		<i>yam</i> ‘stretch’

	...	ṛ	ṛ(C)	...	-an-	-ant-	...
zero	...	<i>kṛ-tá-</i>	<i>dṛṣ-ṭá-</i>	...	<i>rāj-a-su, rāj-ñ-e</i>	<i>cár-at-su</i>	...
full	...	<i>á-kar-at</i>	<i>á-darś-am</i>	...	<i>rāj-an-i</i>	<i>cár-an(t)</i>	...
(Brugmann’s)	...	<i>ca-kār-a</i>	<i>da-dárś-a</i>	...	<i>rāj-ān-am</i>	<i>cár-ant-am</i>	...
long	...	<i>á-kār-ṣam</i>	<i>(a-drāk)</i>	...	–	–	...
		<i>kṛ</i> ‘make’	<i>dṛś</i> ‘see’		<i>-an-</i> nom.suff.	<i>-ant-</i> ptcp.suff.	

**Illustrations**

**A:**

*pat* ‘fall, fly’: 3 sg. redupl. (non-caus.) aor. act. *á-pa-pt-at*, 3 sg. pres. I act. *pát-ati*, 3 sg. caus. pres. act. *pāt-áyati*  
*tap* ‘heat’: 3 sg. pres. I act. *táp-ati*, 3 sg. caus. pres. act. *tāp-áyati*, 3 sg. sigm. aor. act. *a-tāp-sīt*  
*yaj* ‘sacrifice’: ptcp. perf. pass. *iṣ-ṭá-*, 3 sg. pres. I act. *yāj-ati*, 3 sg. sigm. aor. inj. act. *yāṭ*

**Ā:**

*sthā* ‘stand’: ptcp. perf. pass. *sthī-tá-*, 3 sg. root aor. act. *á-sthā-t*, 3 sg. pf. act. *ta-sth-áu* [full grade before a vowel], 3 sg. caus. pres. act. *sthā-p-áyati*

**I:**

*ric* ‘leave’: ptcp. perf. pass. **rik-tá-**, 3 sg. pluperf. act. *á-ri-rec-īt*, 3 sg. perf. act. *ri-réc-a*, 3 sg. sigm. aor. act. **á-raik**

**Ī:**

*nī* ‘lead’: ptcp. perf. pass. **nī-tá-**, 3 sg. pres. I act. **náy-ati**, 3 sg. sigm. aor. subj. act. **ne-ṣ-at**, 2 pl. sigm. aor. inj. act. **nai-ṣ-ṭa**, 3 sg. perf. act. **ni-nāy-a**

**Ṛ:**

*kṛ* ‘make’: ptcp. perf. pass. **kṛ-tá-**, ptcp. root aor. mid. **kr-āṇá-**, 3 pl. root aor. act. **á-kr-an**, 3 sg. them. aor. act. **á-kar-at** AV+, 3 sg. perf. act. **ca-kār-a**, 1 sg. sigm. aor. act. **á-kār-ṣ-am**

*dṛś* ‘see’: ptcp. perf. pass. **dṛṣ-tá-**, 1 sg. root aor. act. **á-darś-am**, 3 sg. perf. act. **da-dárś-a** (3 sg. sigm. aor. act. **a-drāk** Br., **a-drāk-ṣ-īt**)

**Ṝ:**

*tṝ* ‘pass’: ptcp. perf. pass. **tṝ-ṇá-**, 2 sg. pres. VI act. **tṝ-asi**, 3 sg. pres. I act. **túr-ati**, 3 sg. sigm. aor. act. **á-tṝ-īt**

**N:**

**-an-** (nominal stem suffix), stem **rāj-an-** ‘king’: loc. pl. **rāj-a-su**, dat. sg. **rāj-ñ-e**, loc. sg. **rāj-an-i**, acc. sg. **rāj-ān-am**

**-ant-** (act. ptcp. suffix), act. ptcp. **cár-an-** ‘wandering’: loc. pl. **cár-at-su**, nom. sg. **cár-an(t)**, acc. sg. **cár-ant-am**

**Ñ:**

*jan* ‘be born, generate’: ptcp. perf. pass. **jā-tá-**, 3 sg. pres. IV mid. **jā-ya-te**, 3 sg. perf. mid. **ja-jñ-é**, 1 sg. pres. I act. **ján-āmi**, 3 sg. perf. act. **ja-jñn-a**

**Morphophonological classification of (root) morphemes**

In accordance with the ability of the root to appear in all or some of the above-listed ablaut grades, all Vedic roots can be divided into the following three main classes:

- (i) Fully alternating roots (hereafter called “alternating” for short) attest all three morphophonological grades; cf. *ji-/jy-* : *je-/jay-* : *jai-/jāy-*, etc.
- (ii) Non-zeroing roots show the full grade in the forms where the zero grade is expected; cf. *pad* ‘fall’: pres. IV *pád-ya-te* ‘falls’ (instead of \*\*pd-ya-te) etc. In Vedic, a number of roots that in later (Classical) Sanskrit belong to the non-zeroing

type (in particular, most of the *CaC* roots) still attest isolated forms with the zero grade, thus vacillating between the type (i) and type (ii) behavior; cf. the thematized reduplicated present (I–III in the notation of Table 4.18) *pīdamāna-* {*pīpd-*} made from this root.

- (iii) Non-alternating roots appear only either in the full or (for only a few roots) in the long grade; cf. *vīd* ‘be firm’ etc.

Another important division distinguishes between *seṭ* and *aniṭ* roots. *seṭ* roots require insertion of *-i-* before certain suffixes (inf. *-tum*, ptcp. perf. pass. *-tá-*, fut. *-sya-*, sigm. aor. *-s-*); cf. e.g. *vad* ‘say’: ptcp. perf. pass. *ud-i-tá-* (not *\*ut-tá-*) ‘said’; *bhū* ‘become’: inf. *bhāv-i-tum*. Historically, this *i* goes back to the vocalic realization of a laryngeal in the interconsonantal position, but in Sanskrit the class of formations with the inserted *i* was considerably expanded by analogy.

**MORPHOLOGY**

**Nominal morphology**

The grammatical categories of the noun include (i) the case (nominative, vocative, accusative, instrumental, dative, ablative, genitive and locative); (ii) the number (singular, dual and plural); and (iii) the gender (masculine, feminine and neuter).

Alongside inflection proper, the forms of the paradigm may differ in their stems. The two main types of stem alternations include (i) ablaut in the root (only in a few noun stems) or, much more often, in stem suffixes; and (ii) accentuation. The ablauting morpheme may exhibit as many as three grades within the paradigm: (i) long or “Brugmann’s”, (ii) full and (iii) zero; the zero grade may be represented by two variants in some declension types (e.g. in the *an*-declension). Accordingly, in terms of the ablaut type, all declension types can be divided into “three stem”, “two stem” and non-ablauting declensions. Within those types that distinguish between at least two stems, i.e. with the long or full grade, on the one hand, and with the zero grade, on the other, all forms of the paradigm are traditionally divided into “strong” and “weak” forms, respectively. The strong forms include nom., acc. and voc. of the masculine and feminine genders, except for the acc. pl., as well as nom./acc./voc. pl. of the neuter. The weak forms encompass all other members of the paradigm. Among the strong forms, the voc. sg. may differ from the others as the only form that has the full grade (shown with the underline font) as opposed to the long grade of other strong forms. In some declension types, the loc. sg. may have the full grade and, accordingly, be grouped with strong forms.

The endings attested in the declension paradigm are summarized in Table 4.7. The rightmost column shows the pronominal endings that do not occur in the substantive declension.

**TABLE 4.7 OLD INDO-ARYAN CASE ENDINGS**

	<i>Substantive</i>	<i>Pronominal</i>
<i>Singular</i>		
NOM.	<i>-s, -ø</i> (in <i>ā-, ī-, n-, r-</i> declensions)	<i>-ám</i>
VOC.	<i>-ø</i>	
ACC.	<i>-am, -m</i> (in vocalic non-root declensions)	

(Continued)

TABLE 4.7 (CONTINUED)

	<i>Substantive</i>	<i>Pronominal</i>
<i>Singular</i>		
INSTR.	$\bar{a}$ , - $\bar{}$ , - <i>ena</i> (in <i>a</i> -declension, pron. m.)	
DAT.	- <i>e</i> ; - <i>ai</i> (in $\bar{a}$ -, $\bar{i}$ -, $\bar{u}$ -declensions, pron. f.)	- <i>smai</i> (m./n.), - <i>syai</i> (f.)
ABL.	= GEN. (except for <i>a</i> -type); - $\bar{a}t$ (in <i>a</i> -declension)	- <i>smāt</i> (m./n.), - <i>syās</i> (f.)
GEN.	- <i>as</i> , - <i>sya</i> (in <i>a</i> -declension, pron. m.); - $\bar{a}s$ (in $\bar{a}$ -, $\bar{i}$ -, $\bar{u}$ -declensions, pron. f.)	- <i>syās</i> (f.)
LOC.	- <i>i</i> , - $\bar{o}$ ; - $\bar{a}m$ (in $\bar{a}$ -, $\bar{i}$ -, $\bar{u}$ -declensions, pron. f.)	- <i>smin</i> (m./n.), - <i>syām</i> (f.)
NOM./ACC./VOC. <b>n.</b>	- $\bar{o}$ , - <i>m</i> (in <i>a</i> -declension)	- <i>d</i>
<i>Dual</i>		
NOM./ACC./VOC.	$\bar{a}$ , - $\bar{}$ , - <i>au</i>	
INSTR./DAT./ABL.	- <i>bhyām</i>	
GEN./LOC.	- <i>os</i>	
NOM./ACC./VOC. <b>n.</b>	- $\bar{i}$	
<i>Plural</i>		
NOM./VOC.	- <i>as</i>	<i>e</i> (m.) (← . . . <i>a-i</i> )
ACC.	- <i>as</i> , - $\bar{s}$ (in vocalic declensions, f.), - $\bar{n}$ (in vocalic declensions, m.)	
INSTR.	- <i>bhis</i>	
DAT./ABL.	- <i>bhyas</i>	
GEN.	- $\bar{a}m$ , - $\bar{n}\bar{a}m$ (in vocalic declensions)	
LOC.	- <i>su</i>	
NOM./ACC./VOC. <b>n.</b>	- <i>i</i> , - $\bar{}$ , - $\bar{n}\bar{i}$ (in vocalic declensions)	

V = vowel;  $\bar{}$  = length of the stem vowel; boldface = strong forms; underlined = full grade of alternation.

Below are given in table form a few important declension types: simple consonant stems as attested in root nouns; the *nt*-declension (represented, foremost, by present active participles and *mant-/vant-* adjectives); the most productive masculine *a*-type (type *devá-*); and the feminine  $\bar{i}$ -types. The forms unattested in early Vedic but reconstructable and/or known from later texts are in square brackets.

(i) Root consonant stems

((**m./n./f.**)  $^{\circ}k\acute{r}t$ - ‘doing, producing’,  $^{\circ}j\acute{it}$ - ‘winning, conquering’,  $^{\circ}cyút$ - ‘moving’, *soma-sút*- ‘pressing Soma’, *dīrgha-srút*- ‘heard far’); *pád*- ‘foot’, *áp*- ‘water(s)’, (**m./n./f.**) *sáh*- ‘overwinner’, *váh*- ‘carrying, drawing’ (e.g. in *anaḍ-váh*- ‘ox’)

**TABLE 4.8 CONSONANT DECLENSION**

<i>Singular</i>	<i>Alternating nouns</i>			
NOM.	°kʃt	páť		šát
VOC.	°kʃt	[pad]		
ACC.	°kʃt-am	pád-am		sáh-am
INSTR.	°kʃt-ā	pad-á	ap-á	sah-á
DAT.	°kʃt-e	pad-é		sah-é
ABL./GEN.	°kʃt-aḥ	pad-áh	ap-áh	sah-áh
LOC.	°cyút-i	pad-í		
NOM./ACC./VOC. <b>n.</b>	°jít	°pāt <sup>3</sup>		
<i>Dual</i>				
NOM./ACC./VOC.)	°jít-ā, °kʃt-au	pád-ā	áp-ā	sáh-ā °váh-au
INSTR./DAT./ABL.	[. . d-bhyām]	pad-bhyám		
GEN./LOC.	°kʃt-oḥ	pad-óḥ		
NOM./ACC./VOC. <b>n.</b>	[. . t-ī ?]			
<i>Plural</i>				
NOM./VOC.)	°kʃt-aḥ	pád-aḥ	áp-aḥ	sáh-aḥ
ACC.	°kʃt-aḥ	pad-áh	ap-áh	<b>m.</b> sah-áh, <b>mf.</b> sáh-aḥ
INSTR.	somasúd-bhiḥ	pad-bhiḥ	ad-bhiḥ	
DAT./ABL.	°kʃd-bhyaḥ	pad-bhyaḥ	ad-bhyaḥ	šad-bhyaḥ
GEN.	°kʃt-ām	pad-ām		sah-ām
LOC.	°kʃt-su	pat-sú	ap-sú	anaḥ-út-su
NOM./ACC./VOC. <b>n.</b>	(°śrút <sup>4</sup> )			

(ii) (a)nt-declension

(with the root accent: *cárant-* ‘going’ [pres. I *cára-*<sup>ti</sup> ‘go’], *yájant-* ‘sacrificing’ [pres. I *yája-*<sup>ti</sup> ‘sacrifice’], *vádant-* ‘speaking’ [pres. I *váda-*<sup>ti</sup> ‘speak’], *pásyant-* ‘looking’ [pres. IV *pásya-*<sup>ti</sup> ‘look’], *patáyant-* ‘flying’ [AYA-pres. *patáya-*<sup>ti</sup> ‘fly’]; with the suffix accent: *gr̥nánt-* ‘praising’ [pres. IX *gr̥nā-*<sup>ti</sup>, *gr̥nī-*<sup>te</sup> ‘praise’], *jānánt-* ‘knowing’ [pr. IX *jānā-*<sup>ti</sup>, *jānī-*<sup>te</sup> ‘know’], *yánt-* ‘going’ [pres. II *é-*<sup>ti</sup> ‘go’], *sánt-* ‘being’ [pres. II *ás-*<sup>ti</sup> ‘be’]; *paśumánt-* ‘having cattle’)

**TABLE 4.9 (A)NT-DECLENSION**

<i>Singular</i>	“Lengthening” type ( <i>m/vant-</i> )	
NOM.	<i>cáran; jānán</i>	<i>paśu-mán</i>
VOC.	[ <i>cáran</i> ]	
ACC.	<i>cárant-am; gr̥nánt-am</i>	
INSTR.	<i>cárat-ā; jānat-ā</i>	
DAT.	<i>cárat-e; jānat-é</i>	
ABL./GEN.	<i>cárat-aḥ; yát-aḥ</i>	
LOC.	[ <i>cárat-í</i> ]; °yat-í	
NOM./ACC./VOC. <b>n.</b>	<i>cárat; sát</i>	

(Continued)



TABLE 4.9 (CONTINUED)

<i>Dual</i>	
NOM./ACC./VOC.)	<i>cárant-ā, yájant-au; sánt-ā, sánt-au</i> <sup>5</sup>
INSTR./DAT./ABL.	[ <i>cárad-bhyām</i> ]
GEN./LOC.	[ <i>cárat-oḥ</i> ]
NOM./ACC./VOC. n.	[ <i>cárat-ī</i> ]; <i>°yat-ī</i>
<i>Plural</i>	
NOM./VOC.)	<i>pásyant-aḥ</i> <sup>6</sup> ; <i>jánánt-aḥ</i>
ACC.	<i>vádat-aḥ; gr̥nát-aḥ</i>
INSTR.	<i>patáyad-bhiḥ</i>
DAT./ABL.	<i>pásyad-bhyaḥ; yád-bhyaḥ</i>
GEN.	<i>cárat-ām; sat-ām</i>
LOC.	<i>patáyat-su; gr̥nát-su</i>
NOM./ACC./VOC. n.	[ <i>cárant-i</i> ]; <i>sánt-i</i> <sup>7</sup> <span style="float: right;"><i>paśu-mánt-i</i><sup>8</sup></span>

(iii) *a*-stems (*devá-* ‘god’, *yajñá-* ‘sacrifice’, *padá-* ‘foot’, *kárṇa-* ‘ear’)

TABLE 4.10 *A*-DECLENSION (TYPE *DEVÁ-*)

<i>Singular</i>	
NOM.	<i>devá-ḥ</i>
VOC.	<i>déva</i>
ACC.	<i>devá-m</i>
Instr.	<i>devéna, (yajñá)</i>
DAT.	<i>deváya</i>
ABL.	<i>yajñát</i>
GEN.	<i>devásya</i>
LOC.	<i>devé</i>
NOM./ACC./VOC. n.	<i>padá-m</i>
<i>Dual</i>	
NOM./ACC./VOC.)	<i>devá, deváu</i> <sup>9</sup>
INSTR./DAT./ABL.	<i>kárṇābhyām</i>
GEN./LOC.	<i>devá-y-oḥ</i>
NOM./ACC./VOC. n.	<i>padé</i>
<i>Plural</i>	
NOM./VOC.)	<i>devāḥ, devásah</i>
ACC.	<i>devān</i>
INSTR.	<i>devāiḥ, devébhiḥ</i> <sup>10</sup>
DAT./ABL.	<i>devébhyaḥ</i>
GEN.	<i>devānām</i>
LOC.	<i>devéṣu</i>
NOM./ACC./VOC. n.	<i>padā, padāni</i> <sup>11</sup>

(iv) *ī*-declension

**Root stems:** *dhi-* ‘thought’, *śrī-* ‘light, splendour’; *vr̥kī-* type (*{\*-iH-}*, consonant type): *aparī-* ‘future’, *nadī-* ‘river’, *naptī-* ‘grand-daughter’, *rathī-* ‘charioteer’, *vr̥kī-* ‘she-wolf’; *devī-* type (*{-ī-}*, vocalic type): *devī-* ‘goddess’; *māhī-* ‘earth; cow’; *oṣadhī-* ‘herb, plant’

**TABLE 4.11 Ī-DECLENSION (TYPES DEVĪ- AND VṚKĪ-)**

	<i>root stems</i>	<i>vṛkī</i> -type	<i>devī</i> -type
<i>Singular</i>			
NOM.	<i>dhī-h</i>	<i>vṛkī-h</i>	<i>devī</i>
VOC.		<i>°nadi</i>	<i>dēvi</i>
ACC.	<i>dhīy-am</i>	<i>vṛk<sub>i</sub>y-ām</i>	<i>devī-m</i>
INSTR.	<i>dhīy-ā</i>	<i>rath<sub>i</sub>y-ā</i>	<i>devy-ā</i>
DAT.	<i>dhīy-é</i>	<i>vṛk<sub>i</sub>y-è</i>	<i>devy-ái</i>
ABL./GEN.	<i>dhīy-ás</i>	<i>ad<sub>i</sub>y-àh</i>	<i>devy-āḥ</i>
LOC.		<i>nadī</i>	<i>devy-ām</i>
<i>Dual</i>			
NOM./ACC.	<i>°śrīy-ā, -au<sup>(AV+)</sup></i>	<i>nad<sub>i</sub>y-ā</i>	<i>devī</i>
VOC.		<i>rath<sub>i</sub>y-ā</i>	<i>devī</i>
INSTR./DAT./ABL.		[ <i>dhī-bhyām</i> ] [ <i>vṛkī-bhyām</i> ] <i>rōdasī-bhyām</i>	
GEN./LOC.	[ <i>dhīy-óh</i> ]	<i>napt<sub>i</sub>y-òh</i>	<i>rōdas<sub>i</sub>y-óh</i>
<i>Plural</i>			
NOM./ACC.	<i>dhīy-aḥ</i>	<i>rath<sub>i</sub>y-àḥ</i>	<i>devī-h</i>
VOC.	<i>dhīy-aḥ</i>		<i>dēvī-h</i>
INSTR.	<i>dhī-bhiḥ</i>	<i>nadī-bhiḥ</i>	<i>ōśadhī-bhiḥ</i>
DAT./ABL.	[ <i>dhī-bhyáḥ</i> ]	<i>aparī-bhyaḥ</i>	<i>ōśadhī-bhyaḥ</i>
GEN.	<i>dhī-n-ām, (dhīy-ām<sup>1*</sup>)</i>	<i>nadī-n-ām</i>	<i>mahī-n-ām</i>
LOC.	<i>dhī-śu</i>	<i>nadī-śu</i>	<i>ōśadhī-śu</i>

The history of the nominal inflexion is summarized in Table 4.12:

**TABLE 4.12 PIE SOURCES OF CASE ENDINGS**

<i>Singular</i>	
NOM.	-s, -ø < *-s, *-ø
VOC.	-ø < *-ø
ACC.	-(a)m < *-m
INSTR.	-ā, -̄ < *-eh <sub>1</sub> (-ena is borrowed from pronominal declension)
DAT.	-e; -ai < *(e)y (a-declension): -āya < * . . o-ey + ā (particle (?))
ABL.	-āt < *-o-ed (?)
GEN.	-as < *(o)s, -sya (from pronominal) -āyās for *-ās (< * . . eH-os) with insertion of -y- from ī-declension
LOC.	-ī, -ø < *-i, *-ø -ām < *-ā + am ?
NOM./ACC./ VOC. <b>n.</b>	-ø, -m (in a-declension)
<i>Dual</i>	
NOM./ACC./VOC.	-̄ < *-h <sub>1</sub> , -au < ā + u
INSTR./DAT./ABL.	-bhyām < *?
GEN./LOC.	-os < *-h <sub>1</sub> e/oHs ?
NOM./ACC./ VOC. <b>n.</b>	-ī < *-ih <sub>1</sub>
<i>Plural</i>	
NOM./VOC.	-as < *-es -ās < * . . o-es (a-declension)
ACC.	-as < *-ns (< *-m-s (?)) -ān (a-declension): analogy with nom. pl. - n: analogically from a-declension

(Continued)

TABLE 4.12 (CONTINUED)

<i>Plural</i>	
INSTR.	- <i>bhis</i> < *-b <sup>h</sup> i + *-s (plural) - <i>ais</i> (pronominal)
DAT./ABL.	- <i>bhyas</i> : *-b <sup>h</sup> i + *-(y)os (abl. pl.)
GEN.	- <i>ām</i> < *... o-om ( <i>a</i> -declension; then analogically to other types)
LOC.	- <i>su</i> < *-su
NUM./ACC./ VOC. n.	- <i>i</i> , - <sup>-</sup> < *-h <sub>2</sub> ; - <sup>-</sup> <i>n-i</i> is borrowed from <i>n</i> -declension

## Verbal morphology

### *Grammatical categories of the verb*

The Vedic verbal paradigm includes four main classes of forms (“tense systems”), called present, aorist, perfect and (rare in early Vedic) future systems. Within each of these subsets, forms are built on the same stem, i.e. on present, aorist and perfect stems, respectively. There are several sets of personal endings: “primary” (used foremost in the present tense) and “secondary” (endings used in the imperfect, aorist and pluperfect, as well as in two non-indicative moods, the optative and subjunctive), perfect, imperative and subjunctive. Each tense system includes a number of finite forms and a pair of participles, active and middle.

The inventory of the grammatical categories of the verb includes (i) person (1st, 2nd and 3rd) and number (singular, dual and plural); (ii) mood (indicative, imperative, injunctive, subjunctive, optative and conditional); (iii) tense: present, three past tenses (imperfect, perfect and aorist, in Vedic typically encoding a recent past event, most often experienced by the speaker/author of the text) and future; and (iv) diathesis, also called “voice” (active and middle). The range of the functions rendered by the middle diathesis is typical for the ancient Indo-European linguistic type as attested in “Classical” languages (Ancient Greek, Latin). Here belong the self-beneficent meanings with no valency change (‘do smth. for oneself’, as in the textbook example *yájati* ‘sacrifices’ ~ *yájate* ‘sacrifices for oneself’), as well as a number of intransitivizing derivations, such as passive, reflexive and anticausative (decausative); see Gonda 1979. The choice of the function(s) idiosyncratically depends on the base verb. However, already in the language of the earliest text, RV, we observe the loss of several grammatical functions of the ancient Indo-European middle, and the intransitivizing functions are largely taken over by special productive markers, such as the passive suffix -*yá-* and the reflexive pronouns *tanú-* and *ātmán-*; see Kulikov 2012b.

The non-finite forms include two participles (active and middle) for each tense, converbs (traditionally called “absolutives” or “gerunds”), infinitives, gerundives and some others.

### *The morphemic structure of the verbal form*

The verbal form has the following maximal morphemic structure: (preverb(s)) . . . /-(augment *a*-)(reduplication syllable)-root [which may incorporate the nasal infix]- (derivational stem suffix)-(thematic vowel *a*<sup>12</sup>)-(mood)-personal ending. Cf. *ā-cu-cyuv-ī-máhi* (RV 8.9.9) = preverb + reduplication syllable + root + optative morpheme + ending of the 1st person plural middle optative form (= 1 pl. opt. mid. form of the reduplicated (causative) aorist of the verb *cyu* ‘move, impel’) ‘(if) we could impel [you]’; *vṛ-a-dī-dviṣ-a-h* (AV) = preverb + augment + reduplication syllable + root + thematic vowel + secondary (= aorist/imperfect) ending of the 2nd person singular active form (= 2 sg. act. of the reduplicated (causative) aorist of the verb *dviṣ* ‘hate’) ‘you have made (them) hate each other’; {*á-chānd-s-t*} → *áčhān* augment + root + sigmatic aorist morpheme + secondary ending of the 3rd person singular active form (= 3 sg. aor. act. of *chānd* ‘appear’) ‘s/he has appeared’.

*Personal endings and ablaut alternation in the verbal stem*

The general system of personal endings is summarized in Table 4.13. Wherever possible the morphemic border between the ending and the thematic vowel is shown with a hyphen; the lack of a hyphen before the endings of the thematic conjugation shows its coalescence with the thematic vowel (*a*).

**TABLE 4.13 VEDIC VERBAL PERSONAL ENDINGS**

Active diathesis		“primary” (present)		“secondary” (impf., aor., opt.)		imperative		subj.	perf.
		athem.	them.	athem.	them.	athem.	them.		
SG	1	<b>-mi</b>	- <i>mī</i>	<b>-am</b>	- <i>m</i>		(= subj.)	<b>-āni, -ā, -m</b>	<b>-a</b>
	2	<b>-si</b>		<b>-s</b>		- <i>o, -tāt</i>		<b>-s(i)</b>	<b>-tha</b>
	3	<b>-ti</b>		<b>-t</b>		<i>-hi</i> , <sup>13</sup> <i>-dhi</i> , <sup>14</sup> <i>-āna</i> <sup>15</sup>		<b>-t(i)</b>	<b>-a</b>
DU	1	- <i>vás</i>	- <i>vas</i>	- <i>va</i>	- <i>va</i>		(= subj.)	<b>-va</b>	[- <i>vá</i> ]
	2		- <i>thas</i>			- <i>tam</i>		<b>-thas</b>	- <i>áthur</i>
	3		- <i>tas</i>			- <i>tām</i>		<b>-tas</b>	- <i>átur</i>
PL	1	- <i>más(i)</i>	- <i>mas(i)</i>	- <i>ma</i>	- <i>ma</i>		(= subj.)	<b>-ma</b>	- <i>má</i>
	2		- <i>tha, -thana</i>			<i>-ta, -tana</i>		<b>-tha</b>	- <i>á</i>
	3		- <i>nti, -ati</i> <sup>16</sup>	- <i>an, -ur</i> <sup>17</sup>	- <i>n</i>		(- <i>a</i> ) <i>ntu / -atu</i>	<b>-n</b>	- <i>úr</i>

Middle diathesis		“primary”		“secondary”		imperative	subjunctive	perfect
		athem.	them.	athem.	them.			
SG	1		(-) <i>e</i>	- <i>i</i>	- <i>e</i>	(= subj.)	<b>(-)<i>ai</i></b>	- <i>é</i>
	2		- <i>se</i>		- <i>thās</i>	- <i>sva</i>	<b>-se, -sai</b> <sup>18</sup>	- <i>sé</i>
	3		- <i>te (-é)</i>		- <i>ta (-a-t)</i>	- <i>tām, -ām</i>	<b>-te, -tai</b> <sup>19</sup>	- <i>é</i>
DU	1	- <i>vāhe</i>	- <i>vahe</i>	- <i>vahi</i>	- <i>vahi</i>	(= subj.)	<b>-vahi</b>	- <i>vāhe</i>
	2	- <i>āthe</i>	<i>ethe</i>		- <i>āthām / ethām</i>		<b>(-)<i>aithe</i></b>	- <i>āthe</i>
	3	- <i>āte</i>	<i>ete</i>		- <i>ātām / etām</i>		<b>(-)<i>aite</i></b>	- <i>āte</i>
PL	1	- <i>māhe</i>	- <i>mahe</i>	- <i>mahi</i>	- <i>mahi</i>	(= subj.)	<b>-mahe</b>	- <i>māhe</i>
	2		- <i>dhve</i>		- <i>dhvam</i>		<b>-mahai</b>	- <i>dhvé</i>
	3	- <i>āte, -até</i>	- <i>nte</i>	- <i>a(n)ta</i>	- <i>nta</i>	- <i>ātām -ntām</i>	<b>-nta</b>	- <i>ré</i>

The two main classes of stems, athematic and thematic, differ in that the latter appear in the same shape in all forms of the paradigm, while the former exhibit a number of alternations. Like in the declension paradigm, the main types of stem alternations include (i) ablaut in the verbal stems, that is, in the root (in athematic root stems) or in the suffix (in other athematic stems), and (ii) accentuation (partly correlating with the ablaut grade of the morpheme).

The ablauting morpheme may appear (i) in the full (or “Brugmann’s”) grade or (ii) in the zero grade. Accordingly, all forms of the paradigm are divided into “strong” and “weak” forms, respectively. The strong forms (the corresponding endings are in bold-face in the tables) include (i) the sg. act. of the present, imperfect/injunctive and perfect (3 sg. perf. act. has the “Brugmann’s” grade); (ii) 3 sg. imp. act.; (iii) all forms of the

subjunctive; (iv) 3 pl. impf. act. of the reduplicated present (class III) [before *-ur*]. The forms that often or sporadically (shown with the underline font) have the full grade are (iv) 2 pl. pres./imp. act. (almost always before *-tana*, sporadically before *-ta*); and (v) 2 sg. imp. act. (with a few roots). Forms that have the zero grade are called “weak forms”. The accent is on the stem in the strong forms and on the endings in the weak forms (except for some weak forms of the reduplicated present).

### *History of Old Indo-Aryan personal endings*

The rich variety of the Old Indo-Aryan personal endings must go back to the six or seven series of the Proto-Indo-European verbal endings as reconstructed, in particular, by Kortlandt (1979: 66ff.) or Beekes (2011):

**TABLE 4.14 SERIES OF THE PIE VERBAL ENDINGS**

	Primary athematic	Secondary athematic	Primary thematic	Secondary thematic	Perfect	Stative (Middle intr.)	(Middle transitive)
	I	II	IV	III	V	VI	(VII)
SG	1st * <u>-mi</u>	*-m	*-oH	*-om	*-h <sub>2</sub> e	*-h <sub>2</sub>	*-mh <sub>2</sub>
	2nd * <u>-si</u>	*-s	*-eh <sub>1</sub> i	*-es	*-th <sub>2</sub> e	*-th <sub>2</sub> o	*-sth <sub>2</sub> o
	3rd * <u>-ti</u>	*-t	*-e	*-et	*-e	*-o	*-to
PL	1st * <u>-mes</u>	*-me	*-omom	*-omo	*-me	*-med <sup>h</sup> h <sub>2</sub>	*-me(s)d <sup>h</sup> h <sub>2</sub>
	2nd * <u>-th<sub>1</sub>e</u>	*-te	*-eth <sub>1</sub> e	*-ete	*-e	*-d <sup>h</sup> we	*-td <sup>h</sup> we
	3rd *-(e)nti	*-(e)nt	*-o	*-ont	*-(ē)r	*-ro	*-ntro

Using the asterisked Roman numbers to refer to Table 4.14, one may synopsise the history of the attested Vedic personal endings in Table 4.14 (individual endings are explained only where their origin is not straightforwardly explained by the general correspondence given in the heading of the column).

**TABLE 4.15 PIE SOURCES OF VEDIC PERSONAL ENDINGS**

Active diathesis					
SG	1	“primary” < * <b>I</b> ... <i>āmi</i> < *-oH + *-mi (= * <b>I</b> + * <b>IV</b> )	“secondary” < * <b>II</b> <u>-am</u> / -m < *-m̥/-m	imperative (← subj.)	perf. < * <b>V</b>
	2			*-o, *-tōd, *-d <sup>h</sup> i <i>-āna</i> < *... H-ø (+ particle <i>na</i> (?))	
	3			<i>-tu</i> < *-t (* <b>II</b> ) + particle <i>u</i> (← subj.)	
PL	1	<i>-mas(i)</i> < *-mes(-i) (analogically)			
	2	<i>-tha</i> (+ <i>na</i> – particle?)	<i>-ta</i> (+ <i>na</i> – particle?)		
	3	<i>-nti, -ati</i> < *-nti/-ṅti	-(a)n < *(e)nt, <i>-ur</i> < *-r̥ (from * <b>V</b> )		

Middle diathesis

		“primary” < *II + *VI + *I	“secondary” < *VI	imperative	Perfect < “primary” or *VI + *i (from *I)
SG	1	<i>e</i> (analogically)	<i>-i-</i> < * <i>-h<sub>2</sub></i> (*VI)	(← subj.)	
	2	<i>-se</i> (analogically)	<i>-thās</i> (*VI + *II)		
	3	<i>-te</i> < * <i>-t</i> + * <i>-o</i> + * <i>-i</i> (*II + *VI † *I)	<i>-ta</i> (*VII or *II + *VI)	<i>-(t)ām</i> < * <i>-(t)o</i> (*VI/VII)+ particle <i>-am</i>	<i>-é</i> < * <i>-o</i> + * <i>-i</i> (*VI † *I)
PL	1	< *VI + <i>i</i>	<i>-mahi</i> < * <i>-med<sup>h</sup>h<sub>2</sub></i> < *VI + particle (?) <i>m</i>	(← subj.)	
	2				
	3	< * <i>-nt</i> (*II) + * <i>-o</i> (from sg.) + * <i>-i</i> (from act.)	< * <i>-nt</i> (*II) + * <i>-o</i> (from sg.)	< * <i>-nt</i> (*II) + <i>-ām</i> (from sg.)	<i>-ré</i> < * <i>-ro</i> + * <i>-i</i> (*VI † *I)

† = extracted from/under the influence of

*General overview of tense systems and formation of tense stems*

The system of the finite forms of the Vedic Sanskrit verb (and, particularly, its variant attested in the most ancient Vedic text, the RV) is generally considered extremely complicated and irregular as compared to the corresponding system of Classical Sanskrit. Yet this opinion seems to have been imposed by the Sanskritist tradition (essentially going back to the approach of Old Indian grammarians), which usually takes Classical Sanskrit paradigms as a starting point for a grammatical description of Vedic. Such a perspective presents the Vedic paradigms in terms of a list of differences from the Classical Sanskrit system. This approach is, in a sense, anhistorical and methodologically (as well as pedagogically) misleading, since it does not reflect the evolution of the Old Indian morphological system. As is well known, the Classical verbal system evolves from the Vedic, not the other way around. More specifically, the former system can be regarded as a result of reduction of the latter.

To put it differently, the Vedic verbal system shows very few constraints on applying inflectional morphemes to various verbal stems as compared to what we observe in the Classical language. In fact, the Vedic system can be said to be much larger and yet much simpler, in terms of the number of combinatory constraints, as compared to the much smaller Classical system.

The Vedic verbal system can be presented as based on two parameters (see Kulikov 2008): (i) the type of stem and (ii) the type of inflection. There are four types of stems: present (pres.), aorist (acr.), perfect (perf.) and future (fut.). The types of inflection include (1) primary endings (1 sg. act. *-mi*, 2 sg. act. *-si*, 3 sg. act. *-ti*, . . . , 3 pl. act. *-nti/-ati*, etc.); (2) augment *á-* + secondary endings (1 sg. act. *-(a)m*, 2 sg. act. *-s*, 3 sg. act. *-t*, etc.); (3) secondary endings; (4) imperative endings; (5) subjunctive morpheme *a* + subjunctive endings (which are distinct from primary or secondary endings only for some middle forms); (6) optative morpheme *īyā* (which coalesces with the preceding thematic vowel *a* into *e*) + secondary endings; and (7) perfect endings; all these types are exemplified in the upper row of Tables 4.16–17 by the 3 sg. act. and mid. endings. Combining these two sets, we obtain 4 × 7 = 28 logically possible formations. Most of them are actually attested in Vedic (though some are very rare). Only ten of them survive into the Classical

Sanskrit paradigm, including the present and aorist injunctive (which only survives with the prohibitive particle *mā*) and aorist optative (which is only preserved in the precative, based on the root aorist optative). Note that all formations that belong to the standard Classical Sanskrit paradigm (= cells bordered with a shadowed line in Table 4.16) are also present in Vedic, though some of them are very rare (or even exceptional) in the early language, as is the case with the conditional (one attestation in the RV).

Some of these formations require special comments. Thus, “perfects with present endings” include such forms as 3 sg. act. *jāgár-ti* ‘watches’ ( $\sqrt{j}$  ‘become awake’), 3 pl. act. *dīdy-ati* ‘(they) shine’ ( $\sqrt{dī}$  ‘shine’). These formations do not represent a separate synchronic category. Historically, these forms are based on regular perfects, which, at some stage, have been reanalyzed as belonging to the present system – most often on the basis of non-indicative forms (in particular, subjunctives) that share endings with present forms. Once these forms had been transferred to the present system, they lost their synchronic link with the perfect system, becoming present formations with irregular (“perfect”) reduplication (see Kümmel 2000: 59ff., 191–194, 227–230 et passim).

The non-indicative moods of the perfect and aorist are attested only in early Vedic and are non-productive already in the early language (for aorist imperatives, see Baum 2006). Still rarer, even exceptional, are “aorist injunctives with present endings” (labeled by Hoffmann “the type *gathā*”), which include a few isolated forms attested in the RV. Here belong, in particular, 2 du. act. *kṛ-thás* ( $\sqrt{kṛ}$  ‘make’) RV 1.112.8, 5.74.5, 10.39.8; *bhū-thás* ( $\sqrt{bhū}$  ‘become’) RV 6.67.5, 3 du. act. *bhū-tas* RV 10.27.7, and 2 pl. act. *ga-thá* ( $\sqrt{gam}$  ‘go’) RV 8.20.16; 1 pl. med. *dhī-mahe* ( $\sqrt{dhā}$  ‘put’) RV 2.23.10 should perhaps be included in this group, too. These nonce formations (see Lubotsky 1997: 439, 471, 747, 992) do not instantiate root presents (contra Whitney 1885: 21, 34 etc.). As Hoffmann (1952 [1957]: 128–130 [= Hoffmann 1976: 364–366]) has demonstrated, they are built on the basis of root aorist imperatives (such as 2 du. imp. act. *kṛ-tám*, 2 pl. imp. act. *kṛ-ta*, 2 pl. imp. act. *ga-ta* etc.), and on the model of 2 pl. pres. imp. act. *ha-ta* : 2 pl. pres. ind. act. *ha-tha* ( $\sqrt{han}$  ‘kill’). Their status within the inventory of finite forms is clearly marginal.

Future injunctives (= conditionals without augmentations) and optatives are extremely rare. A few such forms are attested in Epic Sanskrit: 2 sg. opt. act. *pari-trāsyah* ( $\sqrt{trā}$  ‘rescue’) MBh 8.13.24, 3 sg. opt. act. *bhaviṣyat* ( $\sqrt{bhū}$ ) MBh 2.51.25, 3 sg. opt. act. *dhakṣyet* ( $\sqrt{dah}$  ‘burn’) MBh 1.221.19 v.l., 3 sg. opt. med. *drakṣyeta* ( $\sqrt{dṛś}$  ‘see’) Rām. 3.1074\*.4; see Whitney 1889: 334; Renou 1930: 401f., 462; Oberlies 2003: 236f., 240.

The system of present formations is the richest subsystem of the verbal paradigm, which is often said to abound in irregular and abnormal forms. The traditional Old Indian system of ten “primary” present classes satisfactorily captures the main morphological types. However, it is also simplistic in some respects, grouping together some types of clearly different nature. In particular, it disregards the important difference between (i) the two main thematic types, classes I (with the accent on the root) and VI (with the accent on the thematic vowel), on the one hand, and (ii) secondary thematizations, i.e. present types that have been transferred from some athematic types to classes I and VI or that can be regarded as such transfers (irrespective of their origin).

Below I outline a somewhat different system of description based on a few parameters that enable us to generate virtually all present formations attested in Vedic. Combining the three main morphological and morphophonological features, we obtain all present stems, including “minor types”, which find no place in the traditional Indian nomenclature. These parameters include (i) the present stem formative (symbolized as F in Table 4.18): no present stem morpheme ( $\emptyset$ ); stem suffixes (*-y(a)-*, *-nó/-nu-* etc.); nasal infix (*-ná/-n-*; placed between  $\cdot \cdot$  in the illustrative examples); and reduplication; (ii) the presence of the thematic vowel *-a-*; and, for thematic present stems, (iii) the placement of the

TABLE 4.16 THE VEDIC VERBAL SYSTEM: EXAMPLES

General survey of tense systems

		Inflection						
		1. Primary endings: -ti, -te	2. Augment + secondary endings: á-, ...-t, á-, ...-ta	3. Secondary endings: -t, -ta	4. Imperative endings: -tu, -tām	5. Subjunctive: a + subj. endings: -a-t(i), -a-te/ta	6. Optative: $\bar{i}/\bar{y}\bar{a}$ + secondary endings: -yā-t/...eṭ, -t-tā/...ēta	7. Perfect endings: -a, -e
PRESENT	Present	Imperfect	(Pres.) injunctive	(Pres.) imperative	(Pres.) subjunctive	(Pres.) optative	Stative (only 3sg., pl. med.)	
AORIST	Aorist with prim. endings (type <i>gathā</i> )	Aorist	Aorist injunctive	Aorist imperative	Aor. subjunctive	Aorist optative (with precative)	<i>cit-e, tītap- e (?)</i>	
PERFECT	(Perfect with prim. [present] endings)	Pluperfect	Perfect injunctive	Perfect imperative	Perfect subjunctive	Perfect optative	Perfect	
FUTURE	Future	Conditional	Conditional without augment	Future imperative	Future subjunctive	Future optative		

Stems

— belongs to the standard Classical Sanskrit paradigm

— (relatively) rare in early Vedic, more common from middle Vedic (Brāhmaṇas) onwards

— rare in early Vedic, exceptional/non-existent after early Vedic

— exceptional in Vedic, non-existent in post-Vedic Sanskrit

— non-existent in Vedic, rare/exceptional in post-Vedic Sanskrit

— non-existent



TABLE 4.17 THE VEDIC VERBAL SYSTEM: EXAMPLES

		Inflection						
	1. Primary endings: -ti, -te	2. Augment + secondary endings: á-...-t, á-...-ta	3. Secondary endings: -t, -ta	4. Imperative endings: -tu, -tām	5. Subjunctive: a + subj. endings: -a-(t), -a-te/tai	6. Optative: $\tilde{i}yā$ + secondary endings: $\tilde{y}ā-t/...et, -\tilde{i}-ta/...eta$	7. Perfect endings: -a, -e	
PRESENT	<i>bhāra-ti, bhāra-te</i>	<i>á-bhara-t, á-bhara-ta</i>	<i>bhāra-t, bhāra-ta</i>	<i>bhāra-tu, bhāra-tām</i>	<i>bhārā-t(i), bhārā-te, -tai</i>	<i>bhāre-t, bhāre-ta</i>	<i>himv-é</i>	
AORIST	<i>2du.act. kṛ-thās, 2pl.act. gā-thā</i>	<i>á-dhā-t, a-dhi-ta</i>	<i>dhā-t, ar-ta</i>	<i>dhā-tu, 2sg.med. kṛ-svā</i>	<i>kār-a-(t), kār-a-te</i>	<i>bhī-yā-t (and prec. bhī-yā-s), ar-t-ta</i>		
PERFECT	<i>jāgár-ti</i>	<i>á-jagan {<sup>o</sup>m-t}</i>	<i>cākān {<sup>o</sup>n-t}</i>	<i>mamāt-tu</i>	<i>cākān-a-t, jijīṣ-a-te, -tai</i>	<i>jagam-yā-t, vavyt-t-ta</i>	<i>cakār-a, cabr-é</i>	
FUTURE	<i>bhaviṣyā-ti, janiṣyā-te</i>	<i>á-bhariṣya-t</i>	<i>bhaviṣya-t</i>	<i>3pl.act. drakṣya-ntu, 2sg.med. tyakṣya-sva</i>	<i>2sg.act. karṣyā-h</i>	<i>dhakṣye-t, drakṣye-ta</i>		

Stems

Roots that occur in examples given in Table 4.17:  $\sqrt{kr}$  'move, raise',  $\sqrt{kan}$  'rejoice',  $\sqrt{kṛ}$  'make',  $\sqrt{gam}$  'go',  $\sqrt{jari}$  'be born, arise',  $\sqrt{ij}$  'become awake',  $\sqrt{yaj}$  'abandon',  $\sqrt{dah}$  'burn',  $\sqrt{dṛṣ}$  'see',  $\sqrt{dhā}$  'put, place',  $\sqrt{bhī}$  'become',  $\sqrt{bhī}$  'carry, bring',  $\sqrt{mad}$  'rejoice, exhilarate',  $\sqrt{vyt}$  'turn',  $\sqrt{hi}$  'impel'

main accent: on the root vs. on the thematic vowel. Note that the stem suffix (or infix) of the secondary thematic classes always appears in the zero grade and contains no vocalic element; for that reason, it cannot bear the accent.

For the sake of convenience, I use a complex notation that is as close as possible to the traditional Indian system of ten classes (symbolized by Roman characters). Each of the secondary thematic types (qualified by the Sanskritist tradition as belonging either to class I or to class VI) is related to the corresponding athematic type (I←V, VI←VII, etc.), thus being presented as the result of thematicization. This is a purely synchronic notation: the arrow (←) does not necessarily mean that the thematic type in question historically goes back to a (hypothetic) athematic pendant. Thus, for instance, I do not argue that, for the class “VI←VII” present *kṛ-n-t-á-ti*, we have to reconstruct the athematic class VII present \**kṛ-ná-t-ti*, etc. A few secondary thematic(ized) types are not actually attested. These include classes I←IX, VI←V<sup>22</sup> and VI←III.

In addition to the nine “primary” classes, this calculus generates one type that is traditionally not included in the system of “primary” present types, passives with the suffix *-yá-* (traditionally considered as one of “secondary” formations, which also include *-áya-*causatives, intensives and desideratives). In fact, the only formal difference between class IV presents and *-yá-*passives is the place of the stress (on the root vs. on the thematic vowel/suffix). Thus, this formal opposition follows the same pattern as the opposition between types I←VII (*śú-m-bh-a-ti*) and VI←VII (*kṛ-n-t-á-ti*). Note that the *-yá-*class also includes a few non-passive *-yá-*presents (symbolized as \*IV in Table 4.18) of the type *mriyáte* ( $\sqrt{mr}$  ‘die’) with passive accentuation (← \**mṛ-ye-te*). On this type, see Kulikov 1997.

There are no athematic presents with the suffix *-i-* in Sanskrit (= athematic counterparts of the *-ya-*presents); one of the few traces of the Proto-Indo-European athematic *i*-type might be the present *kṣéti* ( $\sqrt{kṣi}$  ‘dwell’) < \**tḱ-éy-ti*; see Kortlandt 1989: 109; LIV 2001: 644, note 1.

Next to the main present classes, Table 4.18 also includes two non-productive present types with the suffixes *-cha-* and *-va-*<sup>23</sup> (on which see, in particular, Gotō 1987: 73).

TABLE 4.18 THE VEDIC SYSTEM OF PRESENT STEM TYPES

	present stem formatives (suffixes and other morphemes)								
	ø (no)	<i>-y(a)-</i>	<i>-nó-/nu-</i>	<i>-ná-/nī-</i>	infix <i>-ná-/n-</i>	pres. reduplic.	<i>-ch(a)-</i>	<i>-v(a)-</i>	
athematic: $\sqrt{\square}+F-$	II <i>át-ti</i>	– ( <i>kṣéti</i> )	V, VIII <i>su-nó-ti</i> , <i>ta-nó-ti</i> ( $\sqrt{tan}$ )	IX <i>pr-ná-ti</i>	VII <i>yu-ná-k-ti</i>	III <i>dá-dā-ti</i>	–	(VIII) <i>(tar-u-te)</i>	
thematic: $\sqrt{\square}+F+a-$ accent on the root:	I <i>bháv-a-ti</i>	IV <i>lúbh-ya-ti</i>	I←V <i>i-nv-a-ti</i>	(I←IX) –	I←VII <i>śú-m-bh-a-ti</i>	I←III <i>tí-ṣṭh-a-ti</i>	I <sup>cha</sup> <i>gá-cha-ti</i>	I <sup>va</sup> <i>túr-va-ti</i>	
$\sqrt{\square}+F+a-$ accent on the them. vowel:	VI <i>tud-á-ti</i>	pass., *IV <i>mri-yá-te</i>	(VI←V) <i>(*ṛ-nv-á-ti?)</i>	VI←IX <i>pr-ṇ-á-ti</i>	VI←VII <i>kṛ-n-t-á-ti</i>	(VI←III) –	VI <sup>cha</sup> <i>ṛ-chá-ti</i>	–	

*Voice, causatives and transitivity oppositions*

Voice and passives

There are several verbal formations in Vedic that can be employed in passive constructions. Non-finite passives are passive perfect participles with the suffix *-tá-/ná-* and gerundives, or future passive participles, with the suffixes *-ya-*, *-tavyá-* and *-anīya-*. Finite passive formations include (1) presents with the suffix *-yá-* (derived from the root by means of the suffix *-y(á)-*, which can only take middle endings; e.g. *han* ‘kill’: 1 sg.

*han-yé*, 2 sg. *han-yá-se*, 3 sg. *han-yá-te*, etc.); (2) medio-passive *i*-aorists (with defective paradigm: only 3 sg. in *-i*, 3 pl. in *-ran/-ram* and participle; e.g. *yuj* ‘yoke, join’: 3 sg. *áyoji*, 3 pl. *áyujran*, ptcp. *yujāná-*); (3) middle perfects and statives (which supply passive perfects for some verbal roots; also with defective paradigm: 3 sg. in *-e*, 3 pl. in *-re* and participle; e.g. *hi* ‘impel’: 3 sg. *hinvé* ‘(it) is impelled’, 3 pl. *hinviré* ‘(they) are impelled’; ptcp. *hinvāná-*);<sup>24</sup> and (4) some (isolated) middle forms (e.g. *yuj* ‘yoke, join’: 3 pl. sigm. aor. *ayukṣata*; *gṛ* ‘praise’: 3 sg. class IX pres. *gṛñīté* ‘is praised’).

The inventory of the present passive forms attested in the RV and AV is shown in Table 4.19. The members of the paradigm are mainly exemplified by forms of the verb *yuj* ‘yoke, join’ (which exhibits one of the most complete attested paradigms), supplemented by forms of other verbs where those of *yuj* are unattested. The lacking tense-moods of the passive paradigm (which include imperfect, injunctive, subjunctive and optative) are shown with dark gray shading – with the exception of a few hapaxes marked with middle gray shading.

TABLE 4.19 THE INVENTORY OF THE PRESENT PASSIVE FORMS ATTESTED IN THE RV AND AV

	present	imperfect	injunctive	imperative	subjunctive	optative	
SG	1						
	2	<i>yujyá-se</i>		<i>dhīya-sva</i> AV			
	3	<i>yujyá-te</i>	<i>a-nīya-ta</i> RV <sub>L</sub> <sup>1×</sup>	<i>sīya-ta</i> RV <sub>L</sub> <sup>1×</sup>	<i>dhīya-tām</i> AV	<i>bhriyā-te</i> RV <sup>1×</sup>	
DU	1						
	2						
	3	<i>ucyete</i> RV <sub>L</sub> <sup>1×</sup>					
PL	1	<i>-panyāmahe</i> RV <sub>L</sub> <sup>1×</sup>					
	2			<i>yujya-dhvam</i>			
	3	<i>yujya-nte</i>	<i>-á-sicya-nta</i> AV <sup>1×</sup>	<i>badhya-ntām</i>			
PARTICIPLE		<i>yujyá-māna-</i>					

1× = one attestation; RV<sub>L</sub> stands for late RV.

Only the 3rd person singular and plural forms are well attested. Next to a dozen 2 sg. forms (*yujyáse* ‘you are (being) yoked’, *śasyáse* ‘you are (being) praised’ etc.), we find only one occurrence of a 3 du. form, *ucyete* (RV 10.90.11) ‘[the two feet] are called’, and one (philologically and grammatically rather unclear) form *-panyāmahe*, which may represent a 1 pl. (‘we are (being) glorified’ (?); see Kulikov 2012a: 144–146). 1 sg., 1 du., 2 du. and 2 pl. forms are unattested.

Next to present forms proper, participles and rare imperatives (10 forms or so in the RV and AV), only exceptional attestations of other tense-moods are found. These include as few as five forms:

- (i) 3 sg. impf. *anīyata* ‘(she) was brought’ in the late RV (8.56.4 = Vālahk. 8.4) and 3 pl. impf. *-ásicyanta* ‘(they) were besprinkled’ in AV 14.1.36;
- (ii) 3 sg. inj. *sūyata* ‘(he) is consecrated’ in the late RV (10.132.4) (see Kulikov 2012a: 284ff.) and 3 pl. inj. *-apṛcyanta* in the late RV (1.110.4) ‘(they) united’ (non-pass. intr.) (see Kulikov 2012a: 154);
- (iii) 3 sg. subj. *-bhriyāte* (RV 5.31.12) ‘(it) will be brought’.

Optatives of the present passive do not occur before the very end of the early Vedic period.

Only from the middle Vedic period onwards, when the present passive system becomes well established, do we find a good many imperfects, subjunctives and optatives of -yá-passives.

The early Vedic passive paradigm (for all three tense systems, present, aorist and perfect) is summarized in Table 4.20. An almost complete paradigm is attested for the verbs *su* ‘press (out)’ and *yuj* ‘yoke, join’. (In the cases where forms of these two verbs are unattested, I put in square brackets forms made from other roots.) Different degrees of shading show the status of the corresponding forms: dark gray = lacking and morphologically impossible; middle gray = morphologically possible but unattested or only exceptionally attested (underdeveloped part of the paradigm); light gray = morphologically possible but rare (perhaps primarily for pragmatic reasons; cf. the rarity of passive imperatives).

TABLE 4.20 PASSIVE PARADIGM IN EARLY VEDIC

		present		aorist		perfect/stative	
		indicative	imperative	indicative	inj.	indicative	imper.
SG	1	...	...				
	2	..., <i>yuyáse</i>	[ <i>dhīyasva</i> ]			[ <i>śṛṇviṣé</i> ]	
	3	<i>sūyáte, yuyjáte</i>	[ <i>dhīyátām</i> ]	<i>ásāvi, áyoji</i>	..., <i>yoji</i>	<i>sunvé, yuyujé</i>	[ <i>duháṁ</i> ]
DU	1	...	...				
	2	...	...				
	3	[ <i>ucyete</i> ]	...				
PL	1	[- <i>panyámahe</i> (?)]	...				
	2	...	<i>yuyadhvam</i>				
	3	..., <i>yuyjante</i>	[ <i>badhyantām</i> ]	..., <i>áyujran</i>	...	<i>sunvire, yuyujré</i>	
PARTICIPLE		<i>sūyámāna-, yuyjámāna-</i>		<i>s<sup>h</sup>vāná-, yujāná-</i>		<i>sunvāná-, yuyujāná-</i>	

Furthermore, there are some reasons to assume that stative -āna-participles could have active counterparts, that is, participles derived from the stative stem with the active participle suffix -ant, cf. such forms as *stavánt-* (active participle?) ‘praised’ (first noticed by Watkins 1969: 142ff.) and *pépiśat-* ‘adorned’ (RV 10.127.7; see Schaefer 1994: 45, 152f.), which may point to the unattested stative \**pépiše* ‘is adorned’ of the type *cékite* (Schaefer 1994: 44); (iii) *mahánt-* ‘great’ (probably belonging with the hapax stative *mahe* ‘is able’ (RV 7.97.2) (see Kulikov 2006b: 59ff. for details). These forms represent a good structural parallel to Hittite participles with the suffix -ant- that form passives for transitive verbs (but hitherto have generally been considered isolated phenomena, without parallels in other Indo-European branches), such as *kunant-* ‘killed’ (*kuen-* ‘to kill’) or *dant-* ‘taken’ (*dā-* ‘to take’).

The hypothesized existence of active participles in the stative paradigm, which, in spite of their “active” morphology, were employed in passive constructions, still further supports the connection between the stative formation and the passive syntactic pattern, on the one hand, and, on the other hand, serves as additional evidence against the traditional assumption about a straightforward connection between middle morphology and passive syntax in Vedic.

## Causative oppositions

Vedic attests a remarkable variety of morphological types of causative oppositions and a rich system of morphological causatives.

- (i) The most regular and productive causative marker in the **present** system is the suffix **-(p)áya-** attached to the “Brugmann’s” (full/long) grade of the root,<sup>25</sup> cf. *vṛdh-* ‘grow, increase’ – *vardháyati* ‘makes grow, increases’, *cit-* ‘appear, perceive’ – *ceṭáyati* ‘shows (= makes appear), makes perceive’ (~ *citáyati* ‘appears’).
- (ii) In early Vedic we also find a few other types of the causative opposition within the present system, where the causative member is represented by another present type, usually by one of the presents with **nasal affixes**: suffixes *-nó-/nu-* (pres. class V), *-nā-/nī-* (pres. class IX) or infix *-ná-/n-* (pres. class VII); the intransitive member (anticausative) is often a class IV present (with the suffix *-ya-*) or a class I present. Generally, the intransitive (anticausative) member of the opposition is inflected in the middle, the transitive-causative in the active (but transitives with middle inflection are possible as well). Cf. *i-* ‘go, send’: *éti* (pres. class II) ‘goes’ ~ *inóti*, *invati* (pres. V and its thematization) ‘sends’; *kṣī-* ‘perish, destroy’: *kṣīyate* ‘dies, perishes’ (pres. IV) ~ *kṣināti* (pres. class IX) ‘destroys’; *jan-* ‘be born, arise’: *jāyate* ‘is born’ (pres. class IV) ~ *jānati* (pres. class I) ‘begets’; *pū-* ‘purify’: *pávate* ‘becomes clean, purifies oneself’ (pres. class I) ~ *punāti* (pres. class IX) ‘purifies’.

Two minor types of causative oppositions in the present system include pairs with

- (iii) **reduplicated** causatives (pres. class III, cf. *yúcchati* ‘keeps away’ ~ *yuyóti* ‘makes keep away’) and
- (iv) present **class VI** causatives (cf. *táratī* (pres. class I) ‘crosses’ ~ *tiráti* (pres. class VI) ‘passes over, rescues’).
- (v) Finally, for some verbs, the causative opposition is marked only by the diathesis (**middle: intransitive** ~ **active: transitive**); although some middle forms can also be used transitively (with their self-beneficent meaning), thus showing the labile syntax (Kulikov 2014). This type of opposition is mostly attested for class I presents; see, for instance, Gotō 1987: 48ff., cf.: *námate* ‘bends’ (intr.) ~ *námati* (*/-te*) ‘bends’ (trans.); *váhate* ‘drives, goes’ ~ *váhati* (*váhate*) ‘carries’; *svádate* ‘is sweet’ ~ *svádati* ‘makes sweet’.
- (vi) In the **aoist system**, the causative meaning is typically expressed by the reduplicated aorist, cf. *vṛdh* ‘grow, increase’ – *ávṛydhāt* ‘has made grow’.
- (vii) Furthermore, the intransitive (anticausative) member of the causative opposition can also be expressed by the medio-passive *-i*-aorist, sometimes alongside the sigmatic aorist; cf.: *jan* ‘be born, arise’: *ájani*, sigm. aor. *ájaniṣṭa* ‘has been born’ ~ *ájījanat* ‘has generated’.
- (viii) In the **perfect system**, the causative/non-causative distinction can only be rendered by the diathesis opposition, although not quite consistently: active forms can be used both transitively and intransitively (thus being labile; see Kulikov 2014: 1151–1152). Cf. 3 sg. perf. mid. *paprathé* ‘has spread (intr.)’ ~ 3 sg. perf. act. *paprátha* ‘has spread (trans.), filled out’; 3 sg. perf. mid. *vāvṛdhé*, 3 sg. perf. act. *vavárdha* ‘has grown (intr.)’ ~ 3 sg. perf. act. *vavárdha* ‘has increased (trans.)’ (see, in particular, Kümmel 2000: 319ff., 469ff. et passim).
- (ix) For a number of verbs, there is a correlation between **transitivity and tense**: forms of the present system (i.e. the present proper and imperfect) are employed transitively, while perfect forms are intransitive; cf. 3 sg. perf. *tatána* ‘has stretched’ (intr.) – 3 sg. pres. act. *tanóti*, 3 sg. pres. mid. *tanute* ‘stretches’ (trans.); see Kulikov 1999, where this phenomenon (“split causativity”) is discussed in a typological perspective.

## SYNTAX

Within a short chapter, it is impossible to discuss in detail the main syntactic features of Old Indo-Aryan. Only a few remarkable syntactic peculiarities will be mentioned in this section.

### Passive and ergative(-like) constructions

Passivization typically suggests (i) the promotion of the initial direct object to the subject position (= the subject of the passive construction or passive subject for short) and (ii) the demotion of the initial subject (usually, an agent). The demoted subject either becomes an oblique object (encoded by the instrumental case, as, e.g., in RV 9.86.12d *s<sub>u</sub>vāyudhāḥ soṭṭbhīḥ pūyate vjśā* '[Soma], the well-armed bull, is being purified by pressers'; more rarely by the genitive case (cf. RV 1.61.15a *asmā́ id u tyád ānu dāyy eṣām* 'this very thing was granted to him by them') or, more frequently, remains unexpressed (see Gonda 1951: 77f.), as in RV 9.97.35c *sómaḥ sutáh pūyate ajoyámānaḥ* 'Soma, pressed out, is purified, being anointed.' See Jamison 1979: 133ff.; P. K. Andersen 1986. There are some reasons to treat Vedic constructions with perfect passive participles in *-ta-/-na-* and genitive marking of the agent separately from canonical passives with the instrumental. According to P. K. Andersen (1986), the genitive noun displays a number of subject properties (usually animate; definite and/or refers to old information) in such constructions, and therefore they should be qualified as ergative rather than passive properly speaking.

### Reflexive constructions

The reflexive pronoun *tanú-* has developed from the substantive meaning 'body' and, like its Old Iranian (OAv.) cognate *tanū-*, is well attested in this grammaticalized usage in the RV, as in RV 1.147.2 *vandārus te tan<sub>u</sub>vām vande agne* 'As your praiser, I praise myself, o Agni.' In some cases it is nearly impossible to draw with accuracy the distinction between the reflexive and non-reflexive ('body') meanings: both interpretations are perfectly appropriate in the context, as in RV 10.54.3 *yán mātāraṃ ca pitāraṃ ca sākām / ájanayathās tan<sub>u</sub>vāḥ svāyāḥ* '... since you produced (your) mother and (your) father together from your own body / from yourself.'

Next to the reflexive usages proper, the Vedic reflexive pronouns can be employed in emphatic usages, i.e. as an emphatic reflexive, or intensifier, signaling the fact that its referent is somewhat unexpected in the role where it appears (cf. two usages of Eng. *-self*: *Peter saw himself in the mirror ~ Peter drew this picture himself*); see Kulikov 2007. In the more common adverbial case pattern we find the instrumental forms (as in RV 6.49.13 quoted below). The nominal pattern is attested, for instance, with accusatives and datives, as in AV 1.13.2 = RVKh. 4.4.2 *mṛḍáyā nas tanūbhyo / máyas tokébhyas kṛdhi* 'Be gracious toward ourselves, make pleasure for [our] offspring.'

The reflexive usage of *ātmán-* becomes common after the RV but is still in competition with *tanú-* in the AV (see Kulikov 2007). In Vedic prose, *ātmán* completely ousts *tanú-*; see Delbrück 1888: 207ff., 262f, Wackernagel 1930: 489ff., §240b; and, especially, a brief survey in Oertel 1926, with a rich collection of examples.

The emphatic usage is attested for *ātmán-* from the AV onward, cf. TS 1.7.3.3 *táto devā ábhavan párásurā yásyaivām vidúšo 'nvāhāryā áhriyáte bhávaty ātmánā páráśya bhrātṛvyo bhavati* 'Then the gods prospered, the Asuras perished. He, who, knowing thus, performs the Anvāhārya-rite, prospers himself, his rival perishes.'

In contrast to *ātmán-*, the more archaic stem variant *tmán-* already occurs in emphatic usage in the early RV. Its instrumental appears in the very frequent regular form *tmánā*

(63 attestations in the RV) and in the form *tmányā* (built on the stem *tmānī-* or *tmānya-*, of unclear origin), which occurs in the late RV (1.188.10, 10.110.10) and in the late mantras (VS 20.45 = TB<sup>m</sup> 2.6.8.4 etc.), cf. RV 10.110.10 *upāva sṛja tmányā* ‘Release [the sacrificial animal] yourself.’

### SOME REMARKABLE TYPOLOGICAL FEATURES OF THE LATER (MIDDLE AND NEW) INDO-ARYAN LANGUAGES

Although this chapter mostly concentrates on the earliest chronological layers of Indo-Aryan, which furnish the richest evidence for Indo-European comparative and historical linguistics it will be appropriate to briefly summarize the most salient typological features shared by all or most of the later (Middle and New) Indo-Aryan languages, which continue the developments and result from the innovations and trends starting as early as Old Indo-Aryan.

#### Replacement of the old case system with new agglutinative cases

By the end of the MIA period, that is, at the turn of the second millennium AD, the Indo-Aryan languages lost most of the cases of the original Sanskrit, or Old Indo-Aryan (OIA), system of eight cases<sup>26</sup> (which, except for minor details, is nearly identical to the case system reconstructed for Proto-Indo-European). Generally, only two cases survive, direct (resulting from the merger of nominative and accusative) and oblique (mostly going back to the OIA genitive), although in some languages isolated traces of some other oblique cases, such as the instrumental, locative or ablative, can still be found, sometimes even within the declension paradigm; cf. the Sinhala instrumental case suffix *-en/-in* and Assamese ergative *-e*, both reflecting the OIA instrumental singular ending of *a*-stems *-ena*. The functions of the lost cases are largely taken over by morphemes (bound or free, i.e. postfixes or postpositions) of different origin.

These include:

- (i) Primary, or “old”, postpositions, going back to Proto-Indo-European morphemes that were used in the adpositional function already in the proto-language. An example of an old OIA postposition reflected as a case suffix in a daughter language, in MIA, is the Māhārāṣṭrī ablative suffix *-āhi* < Skr. postposition *adhi* (constructed with the ablative in OIA); see Insler 1991–92; Bubeník 1998: 68f.

Next to old postpositions, there are several markers that result from grammaticalization of some verbal and nominal forms:

- (ii) Postpositions descendant from
  - (ii.a) non-finite verbal forms, in particular, converbs (traditionally called “absolutives” or “gerunds”, cf. Skr. *ādāya* ‘with’, lit. ‘having taken’), gerundives (participia necessitatis) and verbal adjectives;
  - (ii.b) case forms of some nouns (such as Skr. *grhe* ‘in the house’; see below);
- (iii) final members of compounds, which, again, may represent
  - (iii.a) non-finite verbal forms (cf. nominal compounds in *-sthita-*; see below) or
  - (iii.b) nominal case forms (cf. nominal compounds in *-artham* ‘goal, purpose’; see below).

The markers of the first three types (i, ii.a and ii.b), representing free morphemes (words), were originally attached to (non-nominative) case forms of the noun, thus forming, after having become bound morphemes, the second layer of case forms (see below). In type (iii), the source of the new case morpheme was attached to the nominal stem, thus creating a new case within the first layer. In fact, due to the erosion of the nominal inflection by the end of the MIA period, some (oblique) case forms may eventually become indistinguishable from bare stems, and thus the border between types (ii) and (iii) cannot always be drawn with accuracy.

A number of examples of grammaticalization of new postpositions and case suffixes can be found already in the MIA period, in particular, in Apabhraṃśa Prakrits (for details, see Bubeník 1998: 67, 80), cf. the ablative postfix *-t̥hiu* < OIA *sthita-* ‘standing’ (passive perfect participle of the verb *sthā* ‘stand’) and the locative postposition *majjhe* < Skr. *madhye* (loc. sg. of *madhya* ‘middle’) in Apabhraṃśa Prakrit (Bubeník 1998: 67, 80; Bubeník & Hewson 2006: 113):

- a. *hiaya-t̥hiu*  
heart-LOC  
‘out of [my] heart’
- b. *gharaho majjhe*  
house:GEN in  
‘in the house’

In NIA languages we observe a rapid increase in use of such new postpositions, which are normally added to the oblique case form. This grammaticalization may result in the amalgamation of a postposition with the nominal stem or oblique case and, hence, in the rise of a new case. Such is, for instance, the origin of some new case endings, for instance, Sinhala dat. *-ta*, Khovar dat. *-te* < Skr. *-artham* ‘goal, purpose’; Sinhala gen. *-ge* < Skr. *gr̥he* ‘in the house’ (loc. sg. of *gr̥ha-* ‘house’).

Case markers containing *k-* and/or *r-*, which go back to nominal derivatives of the OIA verbal root *kṛ-* (*kar-*) ‘make, do’, can be found in several NIA languages. These include, in particular, genitive morphemes in several NIA languages (see, in particular, Bubeník & Hewson 2006: 122f.), such as Hindi *-kā*, *-ke* < Apabhraṃśa *-kera* < Skr. gerundive *kārya-* ‘to be done’; Awadhi, Maithili *-ker* < Skr. ptcp. perf. pass. *kṛta-* ‘done, made’; Bhojpuri *-kaē* < Skr. adj. *kṛtya-* ‘to be done’.

Likewise, some dative *k-*morphemes, such as Hindi *-ko*, Oriya *-ku*, Marathi *-kē* and Romani *-ke/-ge*, reveal a vestige of the same Sanskrit root *kṛ-* (*kar-*).

The initial stages of the corresponding grammaticalization processes can be dated as early as OIA. Thus, the starting point of the grammaticalization path of Skr. *-artha* ‘goal, purpose’ toward the Sinhala dative case suffix *-ta* is the adverbial usage of the accusative of the Sanskrit bahuvrīhi compounds in *-artha-* (*X-artham*), meaning ‘having X as a goal, purpose’ → ‘for (the sake of) X’: *udakārtham* ‘having water as a goal’ → ‘for water’, *sukhārtham* ‘having happiness, pleasure as a goal’ → ‘for happiness, for pleasure’, *tadartham* ‘having that as a goal’ → ‘for that, therefore’.

In NIA languages, the morphological status of the resulting markers may vary from bound morphemes (case suffixes), tightly connected with the nominal stem (as in Sinhala, cf. examples above), to free morphemes (postpositions). The latter type can be illustrated by the Hindi dative-accusative morpheme *-ko*, which can be shared in some constructions



by several nouns (as in *rām aur mohan ko* ‘to Ram and Mohan’), exemplifying a “Gruppenflexion”, which pleads for a postposition rather than for a suffix analysis.

The difference between these groups of case morphemes is often described in terms of the distinction between cases of the first, second and third layers (Zograf 1976; Masica 1991: 230ff.; Matras 1997). The first layer corresponds to the case in the strict sense of the term and, in Hindi, is limited to the opposition between the direct and oblique cases. The third layer corresponds to clear instances of postpositional phrases, while the second one takes an intermediary position between cases proper and postpositional phrases. It is important to note that only the first-layer case can trigger agreement on adjectives. Although both “Gruppenflexion” and the lack of agreement with second-layer cases appear to distinguish these morphemes from cases proper, the high degree of grammaticalization makes appropriate their association with the category of case in general.

### Productive morphological causatives

NIA languages are notorious for their system of productive morphological causatives, continuing the OIA causatives in *-(p)āya-*, which go back to going back to the Proto-Indo-European *\*-éye/o-*causative. This distinguishes Indo-Aryan from other Indo-European branches and, in particular, from many Western Indo-European languages that lost the remnants of the Proto-Indo-European causatives. From the MIA period onward, we observe the tendency to develop double (second) causatives. Many NIA languages distinguish between contact (direct) and distant (indirect) causatives and may have up to four members of causative chains, differing both in suffixes and root vowel; cf. Hindi *khul-* ‘open’ (intr.) – *khōl-* ‘open’ (trans.), *cal-* ‘go’ – *cal-ā-* ‘drive (make go)’.

### Ergative construction

Many NIA languages exhibit an ergative pattern, usually alongside a nominative-accusative construction (split ergativity). Ergativity is particularly common for Central and Western NIA languages, where ergative pattern appears with perfective verb forms; cf. Punjabi:

*tarkhān-ne kursi-ā banā-ī-ā*  
 carpenter-ERG chair:F-PL.DIR make-PAST.PFV-F.PL  
 ‘The carpenter has made chairs.’

Historically, the ergative construction continues the (late) OIA nominal passive construction with the perfect passive “participle” (more precisely, verbal adjective); see section 5.1.

## FURTHER READING

### Editions and translations of ancient Indo-Aryan texts

Virtually all important Vedic texts have been critically edited, and for most of them there are European translations, with one important lacuna, some parts of the Paippalāda recension of the Atharvaveda (on which see below). The most ancient OIA text, which is also the most important source of information for comparative Indo-European studies, the RV, was edited by Aufrecht (1877), and remains the standard edition; there is also a more

recent edition by van Nooten and Holland (1994), which offers a metrically restored text. There are a number of complete translations of the RV, which include that of Geldner (1951/2003; in German, still remains standard and most often quoted); a new German translation by M. Witzel, T. Gotō and several other scholars (2007–; work in progress; by now two volumes have been published, which encompass books I to V); a Russian translation by Elizarenkova (1989–1999); and the most recent English translation, by Jamison and Brereton (2014). Renou's translation (1955–1969), published in several issues of *Études védiques et pāṇinéennes*, has remained incomplete but nevertheless is of extreme value for a linguistic and philological study of the RV.

For the second most ancient OIA (Vedic) text, the Atharvaveda in the recension Śaunakīya, editio princeps by Shankar Pāndurang Pandit (1895–1898), still preserves its value; the standard European edition by Roth and Whitney (2nd rev. ed. 1924) has no critical apparatus, which is partly reproduced in Whitney and Lanman's (1905) translation. Alongside a number of anthologies, such as Bloomfield 1897, Whitney and Lanman 1905 had remained the only complete translation of the AV (without the last, 20th, book, or kāṇḍa, which almost exclusively is taken from the RV, however) till recently, when Elizarenkova's (2005–2010) Russian translation appeared. Sadly, the second and third volumes of this last work of the great Russian Vedicist were published posthumously, and the last 30 hymns of kāṇḍa XIX have remained untranslated.<sup>27</sup> Another (and presumably more ancient) recension, Paippalāda, which has survived into modern times, became available to Sanskritists in fairly readable Orissa manuscripts only in the middle of the 20th century. The edition, translation and study of the Paippalāda is work in progress, one of the biggest challenges of the contemporary Vedic (and, in general, Indo-Aryan) scholarship. By now, approximately half of the books (kāṇḍas) of this important text have been critically edited and translated; see, in particular, Bhattacharya 1997, 2008 (edition of kāṇḍas I–XVI); Zehnder 1999, Lubotsky 2002, Griffiths 2009 and Kim 2014 (edition and translation of kāṇḍas II, V, VI–VII and VIII–IX, respectively).

Next to numerous printed editions and translations of Vedic texts, there are also a number of online resources, such as the website of Thesaurus Indogermanischer Text und Sprachmaterialien, or TITUS (<http://titus.uni-frankfurt.de/indexe.htm?texte/texte.htm>); GRETEL – Göttingen Register of Electronic Texts in Indian Languages (see <http://gretel.sub.uni-goettingen.de/>); a few cumulative websites providing links to existing resources, such as INDOLOGY (<http://www.indology.info/etexts/>); as well as several resources for MIA and (early) NIA languages, such as, for instance, the Online Pāli Tipiṭaka Website (<http://tipitaka.sutta.org/#/>). Note, however, that some of these online editions may represent mere mechanical transliteration of certain printed editions, not always corresponding to standard scholarly notation,<sup>28</sup> and therefore should be used with caution.

### Historical grammars

Although the main details of the historical grammar of the Indo-Aryan branch are discussed at length in numerous studies and reference books on Indo-European comparative linguistics, a comprehensive updated monographic study tracing the origin of the entire OIA linguistic system from Proto-Indo-Iranian and, further, Proto-Indo-European remains a desideratum. The monograph Burrow 1973 offers an overview of the Sanskrit grammar in a historical perspective, but many of Burrow's analyses and claims (as, for instance, his variant of the laryngeal theory with one single laryngeal phoneme) are questionable and not adopted by most Indo-Europeanists.

Comprehensive historical surveys of the Indo-European origins of the Indo-Aryan phonological systems can be found in Kobayashi 2004 (which covers only the consonant system) and a short overview (chapter) of the origins of the whole phonological system by the same author Kobayashi (forthcoming/2017). An excellent outline of the origin of the OIA morphology from Proto-Indo-Iranian and Proto-Indo-European is Gotō 2013.

The further history of the OIA linguistic system throughout the MIA and NIA periods is outlined in the classical studies Bloch 1934/1965 and Chatterji 1926/1970, which appeared nearly a half century ago and also requires an update. A brief sketch of the linguistic developments from OIA to NIA is given in Oberlies (forthcoming/2017).

### **Synchronic grammatical descriptions and dictionaries of OIA and MIA**

The most comprehensive grammatical description of the OIA linguistic system is the fundamental compendium, consisting of three volumes, started by Jacob Wackernagel at the end of the 19th century with the first volume of his *Altindische Grammatik*, continued after his death by Albert Debrunner (Wackernagel & Debrunner 1896–1957), and still remaining unachieved. The three published volumes encompass phonology, nominal inflection, pronouns, numerals, nominal compounds and nominal suffixes. The fourth volume that should give an exhaustive description of the verbal system is now probably the main desideratum and challenge of Old Indian linguistics. Although a number of monographic studies have appeared by now, dedicated to nearly all verbal formations (that is, individual present and aorist types, the perfect, so-called secondary derivatives and a variety of regular deverbative nominal formations), such as the classical monographs Narten 1964 (on sigmatic aorists), Gotō 1987 (on thematic root presents = Old Indian present class I) and Kümmel 2000 (on perfects), a full systematic overview of the entire verbal system remains a task for future researchers.

There are dozens of more condensed grammars and many reference handbooks on Sanskrit. A few of them have become classics, and have remained such over decades, even though they were written more than a century ago and require updates. These include Whitney 1889 (which encompasses both Vedic and Classical Sanskrit), two Vedic grammars by Macdonell (1910 and 1916) and a very condensed (and therefore hardly appropriate for beginners), yet very rich Vedic grammar by Renou (1952). An extremely well-organized condensed outline of Classical Sanskrit grammar is Zaliznjak 1978 – a masterpiece of grammatical description. For late OIA, as attested in the great Old Indian epics, see Oberlies 2003.

For syntax, Delbrück 1888 still remains probably the most comprehensive monographic study, alongside with two equally old handbooks by Speijer [Speyer] (1886 and 1896). An overview of the most important aspects of the OIA (Vedic) syntax in modern linguistic perspective is given in Kulikov (forthcoming/2017).

The most comprehensive complete dictionary of Sanskrit remains the famous Great Petersburg Dictionary of Sanskrit (Sanskrit-German) compiled by O. Böhtlingk and R. Roth (1855–75) – a true monument of the Sanskrit lexicography. Even in spite of several lacunae due to the fact that several Vedic texts were unavailable (or at least not yet critically edited) at that time, it offers the most complete coverage of the lexical material. Further dictionaries are largely dependent on this chef d'oeuvre, including the most widely used more compact one-volume Sanskrit-English dictionary, Monier-Williams 1899. A new lexicographic project that started in India in the 1970s, *An Encyclopaedic*

*Dictionary of Sanskrit on Historical Principles* (Ghatage et al. 1976–), under the general guidance of A. M. Ghatage (from vol. 3 on, the supervisor of the project was S. D. Joshi; from vol. 8 on, V. P. Bhatta), aims to update and replace the Petersburg Dictionary. This fundamental dictionary is far from completion, however, still dwelling on the character *a*.

There are several concordances of OIA texts. The largest among them, under the general editorship of Vishva Bandhu (1935–1992), covers nearly all Vedic texts, locating all Vedic word-forms attested in the entire Vedic corpus (with relatively few exceptions and lacunae). The concordances for the RV are Lubotsky 1997 (word-forms without translations); Grassmann 1873, out-of-date in many respects but still useful; and Krisch et al. 2006– (ongoing project, aiming to replace Grassmann’s dictionary), both with translations.

Finally, one should mention the grammatical dictionary of the Old Indian (non-denominative) verbs, Whitney 1885, conveniently cumulating all verbal formations derived from each verbal root. This lexicon has several lacunae, and some of Whitney’s interpretations are out-of-date, but nevertheless it remains an indispensable tool for any student of Sanskrit. A dictionary of Old Indian verbs Werba 1997 (only one volume published so far) is very rich and informative, but uses highly cryptic notation and is more difficult for beginners.

For MIA, the most comprehensive overviews are Pischel 1900 (outdated in several parts but still very useful) and Hinüber 2001; see also an important syntactic survey, Bubenik 1998. A convenient didactically oriented survey of individual Middle Indic languages is Mylius 2013, based on the personal teaching material of the author. Grammatical descriptions of individual languages, starting with the oldest documented MIA idiom, Pāli, include important reference grammars, such as Oberlies 2001.

### General and typological overviews of the Indo-Aryan languages

A comprehensive survey of the Indo-Aryan branch is offered in the classic book Masica 1991. Masica 1976 is a pioneering work that defined the major features of the South Asian linguistic area (which includes virtually all Indo-Aryan languages); a short overview of the languages of this area can also be found in Zograf 1990. Zograf 1976 gives an excellent survey of the typologically relevant morphological features of the modern Indo-Aryan languages. Southworth 2005 is a useful overview of the issues arising at the crossroads of linguistics, history and archaeology, intended, above all, for non-linguists, focusing on the aspects of linguistic research that are relevant for those interested in the history and, especially, prehistory of South Asia.

An encyclopedic survey of the major languages of the Indo-Aryan branch is given in the volume edited by G. Cardona and D. Jain (2003) within the *Routledge language family series*. Although this thousand-page compendium is very informative, it is, unfortunately, quite poorly edited, very inconsistent in terminology and inaccurate in glossing,<sup>29</sup> which severely hampers the use of the volume as a reference handbook for the Indo-Aryan languages; besides, it offers rather limited coverage of the Indo-Aryan languages (about 20 languages). Much more comprehensive, methodologically consistent and uniform are two encyclopedic volumes published within the series *Jazyki mira* (Languages of the World), Elizarenkova et al. (2004) (on OIA and MIA) and Oranskaja et al. (2011). The latter volume offers an unprecedented coverage of the NIA languages, containing short but very informative sketches of nearly 40 languages, prepared in accordance with rigorous and methodologically consistent principles.

## ACKNOWLEDGEMENTS

This paper was finalized thanks to a grant from the ERC (grant agreement 313416, EVALISA project) to Jóhanna Barðdal and thanks to a Marie Skłodowska-Curie grant from the European Commission (Grant Proposal number 702895, TRIA project). I would like to thank Alexander Lubotsky for his invaluable remarks, criticisms and comments. I am also grateful to Mate Kapović for careful editing of these chapters.

## NOTES

- 1 This is a somewhat simplified picture, since, in some periods and/or communities, certain MIA languages could even overrun Old Indo-Aryan (Sanskrit) in prestige.
- 2 Except for a few highly infelicitous attempts to introduce the Indian system of accent marking into the Latin transcription, such as the one adopted in Cardona & Jain 2003.
- 3 *dvi-pāt* ‘biped’, *cātuṣ-pāt* ‘quadruped’.
- 4 *dīrgha-śrūt* ‘heard far’ RV 8.25.17.
- 5 Almost exclusively from the late RV on and before vowels: *sánt-au* RV 1.184.1, 10.117.9, *yánt-au* RV 1.139.4.
- 6 *vi-caránt-aḥ* (wrong accent) RV 8.55.4.
- 7 Only *sánti* RV 2.28.1, 8.8.23 (Pādapāṭha *sant-i*).
- 8 *paśumānti* RV 9.92.6, 9.97.1 and *ghṛtāvānti* ‘having ghee’ RV 9.96.13 (Pp. °*ant-i*).
- 9 The ratio of the endings *-ā* : *-āu* in the RV is 7:1.
- 10 *-āiḥ* :: *-ébhiḥ* : RV 1:1 ; AV 5:1 ; Br.+ almost exclusively *-āiḥ*.
- 11 *-ā* :: *-āni* : RV 2:1 ; AV 1:2.
- 12 In the case of thematic and thematicized suffixes such as *-ya-*, *-sa-*, *-nva-* etc., the thematic vowel (*a*) is traditionally regarded as a part of the suffix, the suffixes “properly speaking” being *-y-*, *-s-*, *-nv-*.
- 13 Athematic conjugation, after vowels.
- 14 Athematic conjugation, mostly after consonants.
- 15 Pres. class IX.
- 16 With reduplicated presents (class III).
- 17 With nearly all reduplicated presents (class III), some root presents (class II) and many root aorists.
- 18 Mainly late early and middle Vedic.
- 19 Mainly late early and middle Vedic.
- 20 Only 1× in the RV.
- 21 In the optative.
- 22 An example of this type might be the Ṛgvedic present *ṛṇvati* (\**ṛ-ṇv-á-ti?*) ‘moves, raises’, which does not occur unambiguously accented, however.
- 23 The only formation that might be qualified as the athematic counterpart of this latter type is the RVic hapax *tar-u-te* (*√ṭṛ* ‘pass, overcome’), attested in RV 10.76.2.
- 24 Athematic middle participles with the suffix *-āna-* exhibit unusual syntactic properties in the language of the RV. While the corresponding finite forms are used only transitively, the *-āna-* participles are attested in both transitive and intransitive (passive) patterns (see Delbrück 1888: 264). For instance, the participle *hinvāná-* (root *hi* ‘impel’), taken by all grammars as the middle participle of the nasal present with the suffix *-nó-/nu-* (class V in the Indian tradition), occurs 18 times in intransitive (passive) patterns and 10 times in transitive patterns in the RV). Similarly, the participle *yujāná-* (root *yuj* ‘yoke’) occurs 8 times in intransitive (passive) syntactic patterns (e.g. *rátho yuj-ānāḥ* ‘a chariot that has been yoked’) and 14 times in transitive syntactic patterns (as in *yuj-ānó haritā ráthe* ‘yoking two fallow [horses] to the chariot’). As demonstrated in Kulikov 2006a, 2006b, the grammatical characteristics

of such passive *-āna*-participles should be reconsidered. These participles are grammatically ambiguous; that is, they belong to the following two paradigms: (1) to the paradigm of the (middle) root aorist and medio-passive aorist and (2) to the paradigm of the (middle) present and stative. Thus, the participle *hinvāná* in its transitive use, meaning ‘impelling’, belongs to the paradigm of the transitive nasal present (*hinváte* etc.), but it is a member of the paradigm of the stative (3 sg. *hinvé*, 3 pl. *hinviré*), i.e. a stative participle, when used intransitively (in a passive syntactic pattern), meaning ‘impelled’. Similarly, *yujāná-* is a member of the paradigm of the (transitive) root aorist (*áyukta* etc.) when used transitively (‘yoking’), but it is a member of the paradigm of the passive aorist (3 sg. *áyoji*, 3 pl. *ayujran*), that is, a passive aorist participle, when used in passive syntactic patterns (‘yoked’). Likewise, 3 sg. and 3 pl. middle perfect forms (with the endings *-e* and *-re*, respectively) attested in passive constructions should be taken as statives built on perfect stems, rather than as middle perfects proper. For instance, the form *dadhé* (root *dhā* ‘put’) should be taken to be a 3 sg. form of the middle perfect when it means ‘has put’, and a 3 sg. form of the stative with the passive interpretation ‘is put/has been put’.

- 25 As opposed to the *-áya*-presents with the short grade of the root, which are mostly intransitive.
- 26 For a general survey and discussion of the evolution of the Indo-Aryan case system, see, in particular, Bloch 1934; Zograf 1976; Bubeník 1998: 99–101; Bubeník & Hewson 2006: 102ff.; Masica 1991: 230ff.
- 27 An annotated translation of these hymns is now being prepared by the author of this chapter (Kulikov, in preparation) and will be published as a separate volume, together with a translation of the independent (i.e. not borrowed from the RV) Kuntāpa-hymns of book XX.
- 28 This is, for instance, the case with the TITUS online edition of the Śatapatha-Brahmaṇa, in particular, as far as the word division and accent notation is concerned.
- 29 For instance, readers interested in Indo-Aryan ergativity will find a few remarks on ergative markers and constructions in Hindi and Urdu, but no mention of ergativity in the chapter on Assamese. In fact, however, the authors of this chapter posit two endings of the nominative case, *-e* and *-ø* (p. 419), used in transitive and intransitive constructions, respectively, and thus representing two different cases, ergative and absolutive/nominative. Likewise, readers will not find any mention of converbs (one of the remarkable typological features of Indo-Aryan languages); instead, individual chapters use a few less standard terms such as “conjunctive participle” or “absolutive”. Furthermore, the chapter on the earliest attested Indo-Aryan language, Sanskrit, uses heavily Pāṇini-dependent notation and is barely accessible for non-Sanskritists. These are obvious editorial slips and inconsistencies; see detailed criticism in the reviews Kulikov 2004 and Peterson 2006.

## REFERENCES

### Editions and translations of early Vedic texts (Ṛgveda, Atharvaveda)

- Aufrecht, Thomas 1877, *Die Hymnen des Rigveda, Herausg. Von Th. Aufrecht, 2 Bde. 2. Aufl.*, Bonn
- Bhattacharya, Dipak 1997, *The Paippalāda Saṃhitā of the Atharvaveda: Critically ed. by [ . . . ] Dipak Bhattacharya. Vol. One: Consisting of the first fifteen Kāṇḍas [= books I–XV], (Bibliotheca Indica; 318), The Asiatic Society, Calcutta*
- Bhattacharya, Dipak 2008, *The Paippalāda Saṃhitā of the Atharvaveda: Critically ed. by [ . . . ] Dipak Bhattacharya. Vol. Two: Consisting of the Sixteenth Kāṇḍa [= book XVI], (Bibliotheca Indica; 319), The Asiatic Society, Calcutta*

- Bloomfield, Maurice 1897, *Hymns of the Atharva-veda Together with Extracts from the Ritual Books and Commentaries* (Sacred Books of the East; 42), Clarendon Press, Oxford
- Elizarenkova, Tat'jana Ja. 1989–1999, *Rigveda*, 3 vols. [Russian], Nauka, Moskva
- Elizarenkova, Tat'jana Ja. 2005–2010, *Atharvaveda* (Šaunaka), 3 vols. [Russian] Vost. lit., Moskva
- Geldner, Karl Friedrich 1951, *Der Rig-Veda aus dem Sanskrit ins Deutsche übersetzt . . .* Bd. 1–3 (Harvard Oriental Series; 33–35), Harvard University Press, Cambridge, MA [Repr. in one volume (Paperback) as Harvard Oriental Series; 63 (2003)]
- Griffiths, Arlo 2009, *The Paippalādasamhitā of the Atharvaveda. Kāṇḍas 6 and 7: A New Edition with Translation and Commentary* (Groningen Oriental Studies; 22), Forsten, Groningen
- Jamison, Stephanie W. & Brereton, Joel P. 2014, *The Rigveda : The Earliest Religious Poetry of India*, transl. by Stephanie W. Jamison and Joel P. Brereton, Vol. 1–3, Oxford University Press, New York
- Kim, Jeong-Soo 2014, *Die Paippalādasamhitā des Atharvaveda: Kāṇḍa 8 und 9; eine neue Edition mit Übersetzung und Kommentar*, Röll, Dettelbach
- Kulikov, Leonid (in preparation), *Atharvaveda* (Šaunaka). Pervod s vedijskogo jazyka i komm. L. I. Kulikova, T. 4: Knigi XIX.43–73 i XX (kuntāpa), Vost.lit., Moskva
- Lelli, Ducio 2015, *The Paippalādasamhitā of the Atharvaveda, Kāṇḍa 15: A New Edition with Translation and Commentary*, Leiden University [PhD dissertation]
- Lubotsky, Alexander M. 2002, *Atharvaveda-Paippalāda. Kāṇḍa five. Text, translation, commentary* (Harvard Oriental Series; Opera Minora 4), Harvard University Press, Cambridge, MA
- Nooten, Barend A. van & Holland, Gary B. 1994, *Rig Veda: A Metrically Restored Text with an Introduction and Notes* (Harvard Oriental Series; 50), Harvard University Press, Cambridge, MA
- Pandit, S.P. 1895–1898, *Atharvaveda Samhitā. With the commentary of Sāyanāchārya*, Ed. by Shankar Pāndurang Pandit, Government Central Book Depôt, Bombay
- Renou, Louis 1955–1969, *Études védiques et pāninéennes*, Vol. 1–17, Boccard, Paris
- Roth, Rudolf & Whitney, William Dwight 1924, *Atharva Veda Samhita*, Herausgegeben von R. Roth und W.D. Whitney. 2. verbesserte Auflage besorgt von Max Lindenau, Berlin. [books I–XIX]
- Whitney, William Dwight & Lanman, Charles R. 1905, *Atharva-Veda Samhitā*. Translated into English with Critical Notes and Exegetical Commentary by William Dwight Whitney . . . Revised and edited by Charles Rockwell Lanman (Harvard Oriental Series; 7–8), Harvard University Press, Cambridge, MA
- Witzel, Michael et al. 2007, *Rig-Veda: das heilige Wissen. Erster und zweiter Liederkreis* / aus dem vedischen Sanskrit übers. und hrsg. von Michael Witzel und Toshifumi Gotō. Unter Mitarbeit von Eijirō Dōyama und Mislav Ježić, Verlag der Weltreligionen, Frankfurt/Main–Leipzig
- Witzel, Michael et al. 2013, *Rig-Veda: das heilige Wissen. Dritter bis fünfter Liederkreis* / aus dem vedischen Sanskrit übers. und hrsg. von Michael Witzel (Buch III), Toshifumi Gotō (Buch IV) und Salvatore Scarlata (Buch V), Verlag der Weltreligionen, Berlin
- Zehnder, Thomas 1999, *Atharvaveda-Paippalāda: eine Sammlung altindischer Zaubersprüche vom Beginn des 1. Jahrtausends v. Chr. Buch 2. Text, Übersetzung, Kommentar* (Wissenschaftliche Schriften: Reihe 3. Beiträge zur Sprach- und Literaturwissenschaft, 107), Schulz-Kirchner Verlag, Idstein

### Historical grammars

- Bloch, Jules 1934, *L'indo-aryen: du véda aux temps modernes*, Adrien-Maisonneuve, Paris [Engl. revised ed.: Bloch, J. 1965, *Indo-Aryan from the Vedas to modern times*, English ed. largely rev. by the author and transl. by Alfred Master, Adrien-Maisonneuve, Paris]
- Burrow, Thomas 1973, *The Sanskrit Language*, [3rd] rev. ed., Faber and Faber, London
- Chatterji, S.K. 1926, *The Origin and Development of the Bengali Language*, Allen & Unwin, London [2nd ed.: Allen & Unwin, London 1970]
- Gotō, Toshifumi 2013, *Old Indo-Aryan Morphology and its Indo-Iranian Background, in co-operation with J.S. Klein and V. Sadovski*, Verlag der Österreichischen Akademie der Wissenschaften, Wien
- Kobayashi, Masato 2004, *Historical Phonology of Old Indo-Aryan Consonants*, ILCAA, Tokyo University of Foreign Studies, Tokyo
- Kobayashi, Masato forthcoming/2017, "Phonology of Indic," in: Klein, Jared, Joseph, Brian & Fritz, Matthias (eds.), *Handbook of Comparative and Historical Indo-European Linguistics*, (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK); 41), Walter de Gruyter, Berlin
- Oberlies, Thomas forthcoming/2017, "The Evolution of Indic," in: Klein, Jared; Joseph, Brian & Fritz, Matthias (eds.), *Handbook of Comparative and Historical Indo-European Linguistics*, (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK); 41), Walter de Gruyter, Berlin
- Watkins, Calvert 1969, *Indogermanische Grammatik*. Bd III: *Formenlehre*. 1. Teil. *Geschichte der indogermanischen Verbalflexion*, Winter, Heidelberg

### Handbooks and grammatical studies of Vedic and Classical Sanskrit

- Andersen, Paul Kent 1986, "The Genitive Agent in Rigvedic Passive Constructions," in: Slawski, F. (ed.), *Collectanea linguistica in honorem Adami Heinz*, Zakład Narodowy im. Ossolińskich, Wrocław: 9–13
- Baum, Daniel 2006, *The Imperative in the Rigveda*, LOT, Utrecht
- Delbrück, Berthold 1888, *Altindische Syntax*, Verlag der Buchhandlung des Waisenhauses, Halle a.S. (B. Delbrück. Syntaktische Forschungen. V)
- Gonda, Jan 1951, *Remarks on the Sanskrit Passive* (Orientalia Rheno-Traiectina; 4), Brill, Leiden
- Gonda, Jan 1979, *The Medium in the R̥gveda* (Orientalia Rheno-Traiectina; 24), Brill, Leiden
- Gotō, Toshifumi 1987, *Die "I. Präsensklasse" im Vedischen: Untersuchung der vollstufigen thematischen Wurzelpräsentia*, (SbÖAW 489; VKSKS 18), Verlag der Österreichischen Akademie der Wissenschaften, Wien [2., überarbeitete und ergänzte Aufl. 1996]
- Hock, Hans Henrich 2016, "Pāṇini's language: real or not?" in: Byrd, Andrew Miles; DeLisi, Jessica & Wenhe, Mark (eds), *Tavet Tat Satyam: Studies in honor of Jared S. Klein on the occasion of his seventieth birthday*, Beech Stave Press, Ann Arbor–New York: 101–112
- Hoffmann, Karl 1952 [1957], "Zur vedischen Verbalflexion," *Münchener Studien zur Sprachwissenschaft* 2: 121–137 [= Hoffmann, K. *Aufsätze zur Indoiranistik*, Bd. 2, Reichert, Wiesbaden 1976: 358–372]
- Jamison, Stephanie W. 1979, "The Case of the Agent in Indo-European," *Die Sprache* 25: 129–143



- Kulikov, Leonid 1997, "Vedic *mriyáte* and Other Pseudo-passives: Notes on an Accent Shift," in: Hegedűs, Irén et al. (eds), *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V. Shevoroshkin* (JIES monograph series; 22), Institute for the Study of Man, Washington D.C.: 198–205
- Kulikov, Leonid 2006a, "Passive and Middle in Indo-European: Reconstructing the Early Vedic Passive Paradigm," in: Abraham, Werner & Leisiö, Larisa (eds), *Passivization and Typology: Form and Function* (Typological studies in language; 68), Benjamins, Amsterdam: 62–81
- Kulikov, Leonid 2006b, "The Vedic Medio-Passive Aorists, Statives and Their Participles: Reconsidering the Paradigm," in: Tikkanen, Bertil & Hettrich, Heinrich (eds), *Themes and Tasks in Old and Middle Indo-Aryan Linguistics. Papers of the 12th World Sanskrit Conference, Vol. 5*, Motilal Banarsidass, Delhi: 45–63
- Kulikov, Leonid 2007, "The reflexive pronouns in Vedic: A diachronic and typological perspective," *Lingua* 117/8: 1412–1433
- Kulikov, Leonid 2008, "Two Calculi for a Description of the Vedic Verbal System," in: Arxipov, A. V.; Fedorova, O. V. et al. (eds), *Fonetika i nefonetika. K 70-letiju Sandro V. Kodzasova* [Phonetics and non-phonetics. Felicitation volume on the occasion of the 70th birthday of Sandro V. Kodzasov], *Jazyki slavjanskix kul'tur*, Moskva: 579–587
- Kulikov, Leonid 2012a, *The Vedic -ya-Presents: Passives and Intransitivity in Old Indo-Aryan* (Leiden Studies in Indo-European; 19), Rodopi, Amsterdam
- Kulikov, Leonid 2012b, "Voice and valency derivations in Old Indo-Aryan in a diachronic typological perspective: the degrammaticalization of the middle and other trends in the Vedic verbal system," in: Melazzo, Lucio (ed.), *Usare il presente per spiegare il passato. Teorie linguistiche contemporanee e lingue storiche. Atti del XXXIII Convegno della Società Italiana di Glottologia*. Palermo, 16–18 ottobre 2008 (Biblioteca della Società Italiana di Glottologia; 35), Il Calamo, Roma: 161–191
- Kulikov, Leonid 2013, "Language vs. grammatical tradition in Ancient India: How real was Pāṇinian Sanskrit? Evidence from the history of late Sanskrit passives and pseudo-passives," *Folia Linguistica Historica* 34: 59–91
- Kulikov, Leonid 2014, "The decline of labile syntax in Old Indo-Aryan: A diachronic typological perspective," *Linguistics* 52/4 (special issue: *Typology of labile verbs: Focus on diachrony*): 1139–1165
- Kulikov, Leonid forthcoming/2017, "The Syntax of Indic," in: Klein, Jared; Joseph, Brian & Fritz, Matthias (eds), *Handbook of Comparative and Historical Indo-European Linguistics* (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK); 41), Walter de Gruyter, Berlin
- Kümmel, Martin 2000, *Das Perfekt im Indoiranischen*, Reichert, Wiesbaden
- Macdonell, Arthur Anthony 1910, *Vedic Grammar*, Trübner, Strassburg
- Macdonell, Arthur Anthony 1916, *A Vedic Grammar for Students*, Clarendon Press, Oxford
- Narten, Johanna 1964, *Die sigmatischen Aoriste im Veda*, Harrassowitz, Wiesbaden
- Oberlies, Thomas 2003, *A Grammar of Epic Sanskrit*, (Indian Philology and South Asian Studies; 5), de Gruyter, Berlin–New York
- Oertel, Hanns 1926, *The Syntax of Cases in the Narrative and Descriptive Prose of the Brāhmaṇas. I. The Disjunct Use of Cases*, Winter, Heidelberg
- Renou, Louis 1930, *Grammaire sanscrite*, Adrien-Maisonneuve, Paris (2me éd. revue, corrigée et augmentée: 1960)
- Renou, Louis 1952, *Grammaire de la langue védique*, Lyon
- Schaefer, Christiane 1994, *Das Intensivum im Vedischen*, Vandenhoeck und Ruprecht (HS; Ergänzungsheft 37), Göttingen
- Speijer [Speyer], Jacob S. 1886, *Sanskrit Syntax*, Brill, Leyden

- Speijer [Speyer], Jacob S. 1896, *Vedische und Sanskrit-Syntax*, (Grundriss der Indo-Arischen Philologie und Altertumskunde; Bd. I, Heft 6), Trübner, Strassburg
- Wackernagel, Jacob & Debrunner, Albert 1896–1957, *Altindische Grammatik*, 3 Bde. Vandenhoeck & Ruprecht, Göttingen
- (1896) J. Wackernagel, Bd. I, Lautlehre
- (1905) J. Wackernagel, Bd. II/1, Einleitung zur Wortlehre. Nominalkomposition
- (1930) J. Wackernagel, Bd. III, Nominalflexion – Zahlwort – Pronomen
- (1954) A. Debrunner, Bd. II/2, Die Nominalsuffixe
- (1957) Introduction générale par L. Renou
- (1957) A. Debrunner, Nachträge zu Bd. I
- Whitney, William Dwight 1889, *Sanskrit Grammar*, Harvard University Press, Cambridge
- Zaliznjak, Andrej A. 1975, “Morfonologičeskaja klassifikacija drevneindijskix glagol’nyx kornej [A morphophonological classification of the Old Indian verbal roots],” in: Elizarenkova, Tat’jana Ja. (ed.), *Očerki po fonologii vostočnyx jazykov* [Essays on the phonology of oriental languages], Nauka, Moskva: 59–85
- Zaliznjak, Andrej A. 1978, “Grammatičeskij očerk sanskrita [A grammatical sketch of Sanskrit],” in: Kočergina, Vera A. (ed.), *Sanskritsko-russkij slovar’*, Russkij jazyk, Moskva: 785–895 [2nd ed. 1987]

### Grammars and handbooks of Middle Indic

- Bubeník, Vít 1998, *A Historical Syntax of Late Middle Indo-Aryan* (Apabhraṃśa) (Current issues in linguistic theory; 165), Benjamins, Amsterdam
- Hinüber, Oskar von 2001, *Das ältere Mittellindisch im Überblick*, 2., erw. Aufl., Österreichische Akademie der Wissenschaften, Wien
- Mylius, Klaus 2013, *Zur Didaktik mittelindischer Sprachen* (Beiträge zur Kenntnis südasiatischer Sprachen und Literaturen; 23), Harrassowitz, Wiesbaden
- Oberlies, Thomas 2001, *Pāli – A Grammar of the Language of the Theravāda Tipiṭaka*, (Indian Philology and South Asian Studies; 3), de Gruyter, Berlin
- Pischel, Richard 1900, *Grammatik der Prakrit-Sprachen*, (Grundriss der indo-arischen Philologie und Altertumskunde; 1/8), Trübner, Strassburg

### Dictionaries, grammatical dictionaries, concordances

- Böhtlingk, Otto & Roth, Rudolf 1855–1875, *Sanskrit-Wörterbuch*, Eggers, St. Petersburg
- Ghatage, A.M., [Joshi, S.D.; Bhatta, V.P.] et al. (eds.) 1976–, *An Encyclopaedic Dictionary of Sanskrit on Historical Principles*, Deccan College Postgraduate and Research Institute, Poona
- Grassmann, Hermann 1873, *Wörterbuch zum Rig-Veda*, Brockhaus, Leipzig
- Krisch, Thomas [et al.] 2006–, *RIVELEX: Rigveda-Lexikon*, Bd. 1 (2006); Bd. 2 (2012) (Grazer vergleichende Arbeiten; 20, 26), Leykam, Graz
- Lubotsky, Alexander M. 1997, *A Rgvedic Word Concordance, 2 Parts* (American Oriental Series; 82), American Oriental Society, New Haven
- Monier-Williams, Monier 1899, *A Sanskrit-English Dictionary*, Clarendon Press, Oxford
- Vishva Bandhu (general ed.) 1935–1992, *A Vedic Word-Concordance. I. Samhitās*. 6 vols.; II. The Brāhmaṇas and the Āraṇyakas. 2 vols.; III. Upaniṣads. 2 vols.; IV. Vedāṅga-sūtras, 4 vols, Vishveshvaranand Vedic Research Institute, Lahore–Hoshiarpur
- Werba, Chlodwig H. 1997, *Verba Indoarica: die primären und sekundären Wurzeln der Sanskrit-Sprache, Pars I: Radices Primariae*, Verlag der Österreichischen Akademie der Wissenschaften, Wien
- Whitney, William Dwight 1885, *The Roots, Verb-forms, and Primary Derivatives of the Sanskrit Language*, Breitkopf and Härtel, Leipzig

### Overviews of Indo-Aryan languages

- Cardona, George & Jain, Dhanesh (eds.) 2003, *The Indo-Aryan Languages* (Routledge Language Family Series; 2), Routledge, London–New York [see reviews Kulikov 2004 and Peterson 2006]
- Elizarenkova, Tatjana, Kibrik, Andrej & Kulikov, Leonid (eds.) 2004, *Jazyki mira: Indoarijskie jazyki drevnego i srednego periodov* [Languages of the World: Old and Middle Indo-Aryan Languages], Academia, Moscow
- Kulikov, Leonid 2004, Review of: Cardona & Jain (eds.) 2003, *Acta Orientalia Vilnensia* 5: 286–292
- Masica, Colin P. 1976, *Defining a Linguistic Area: South Asia*, University of Chicago Press, Chicago
- Masica, Colin P. 1991, *The Indo-Aryan Languages*, Cambridge University Press, Cambridge
- Oranskaja, Tat'jana I., Mazurova, Julia V., Kibrik, Andrej A., Kulikov, Leonid I. & Rusakov, Aleksandr Ju. (eds.) 2011, *Jazyki mira: Novye indoarijskie jazyki* [Languages of the World: New Indo-Aryan Languages], Academia, Moscow
- Peterson, John 2006, Review of: Cardona & Jain (eds.) 2003, *Language* 82/4: 891–894
- Southworth, Franklin C. 2005, *Linguistic Archaeology of South Asia*, RoutledgeCurzon, London
- Zograf, Georgij A. 1976, *Morfologičeskij stroj novyx indoarijskix jazykov: opyt strukturno-tipologičeskogo analiza* [The Morphological Structure of the New Indo-Aryan Languages: A Structural-Typological Approach], Nauka, Moskva
- Zograf, Georgij A. 1990, *Jazyki južnoj Azii* [Languages of South Asia], Nauka, Moskva

### Other works

- Beekes, Robert S.P. 2011<sup>2</sup> [Revised and corrected by Michiel de Vaan], *Comparative Indo-European Linguistics*, John Benjamins, Amsterdam–Philadelphia
- Bubenik, Vit & Hewson, John 2006, From case to adposition: *The development of configurational syntax in Indo-European languages* [Current issues in linguistic theory 280], Benjamins, Amsterdam
- Inslar, Stanley 1991–92, “The Prakrit ablative in -āhi,” *Annals of the Bhandarkar Oriental Research Institute* 72-73: 15–21
- Kortlandt, Frederik H. H. 1979, “Toward a reconstruction of the Balto-Slavic verbal system,” *Lingua* 49: 51–70
- Kortlandt, Frederik H. H. 1989, “Lithuanian statýti and Related Formations,” *Baltistica* 25/2: 104–112
- Kulikov, Leonid 1999, “Split causativity: remarks on correlations between transitivity, aspect, and tense,” in: Abraham, Werner & Kulikov, Leonid (eds), *Tense-aspect, transitivity and causativity. Essays in honour of Vladimir Nedjalkov* (Studies in Language Companion Series; 50), Benjamins, Amsterdam: 21–42
- LIV = *Lexikon der Indogermanischen Verben*, Rix, Helmut (ed.), Dr. Ludwig Reichert Verlag, Wiesbaden 2001<sup>2</sup>
- Matras, Yaron 1997, “The typology of case relations and case layer distribution in Romani,” in: Matras, Y.; Bakker, P. & Kyuchukov, H. (eds), *The Typology and Dialectology of Romani*, Benjamins, Amsterdam: 61–93
- Witzel, Michael 1995, “Early Indian culture: Linguistic and textual parameters,” in: Erdosy, G. (ed.). *The Indo-Aryans of ancient South Asia: language, material culture and ethnicity* (Indian philology and South Asian studies; 1), de Gruyter, Berlin: 85–125

# IRANIAN

*Nicholas Sims-Williams*

## INTRODUCTION

At present, Indo-European languages of the Iranian group are spoken over a wide area including virtually the whole of Iran, Afghanistan and Tajikistan, together with the neighbouring parts of Turkey, Syria, Iraq, Pakistan and Uzbekistan, as well as larger or smaller enclaves in Oman, Armenia, Georgia, Azerbaijan, Turkmenistan and western China (see the map at the end of Schmitt 1989). In mediaeval times Iranian languages such as Sogdian and Khotanese were well established even further east, in the area which later became Chinese Turkestan (Xinjiang); at a still earlier period, the original homeland of the Iranian-speaking peoples seems to have lain to the northeast of the present state of Iran.

The present chapter will concentrate on the earliest attested Iranian languages, Avestan and Old Persian, which are naturally the most important for Indo-European studies. These two Old Iranian languages will be described to a large extent in terms of their similarities to and differences from the closely related Old Indian, a procedure justified in the first place by pragmatic considerations. Old Indian (Vedic and Sanskrit) is attested by a huge and varied corpus of literature, written in a clear, almost phonemic script which allows the phonological and morphological structure of the language to be clearly perceived. On the other hand, each of the attested Old Iranian languages is known from a limited corpus – in the case of Old Persian, a tiny corpus – of rather repetitive and monotonous texts, one written in an ambiguous cuneiform writing system, the other by means of an over-elaborate, almost phonetic alphabet, whose intricacies obscure rather than illuminate its grammatical structure. Although some of these deficiencies are made good by the more abundant Middle and Modern Iranian material, it cannot be denied that Iranian evidence is usually more difficult than Indian for a student of Indo-European to evaluate.

From a theoretical point of view, too, it is proper to treat the Iranian languages in constant comparison with Indian, since the two groups are not merely closely related but jointly constitute a single Indo-Iranian branch of the Indo-European family, as is indicated by the innumerable phonological, morphological and lexical isoglosses which they share to the exclusion of all other branches of Indo-European. One such isogloss is the use of OInd. *ārya-*, Av. *a'riia-*, OPers. *ariya-* (from which the name of the country “Iran” derives) as a self-designation for the speakers of Indo-Iranian, whence the alternative term “Aryan”. The closeness of the relationship between Indian and Iranian is most clearly demonstrated by the fact that it is possible to find not just words but whole sentences in Vedic or Avestan which may be transposed from the one language into the other merely by observing the appropriate phonological rules; e.g. Av. *təm amauuaṇtəm yazatəm sūrəm dāmōhu səuuištəm miθrəm yazāi* ‘this powerful, strong (being) worthy of worship, Mithra, the strongest amongst creatures, I shall worship’ (Yasht 10.6) = Ved. *\*tām āmavantam yajatām śūran dhāmasu śāviṣṭham mitrām yajai* (Jackson 1892: xxxi–xxxii).

Despite the overwhelming similarity of Indian and Iranian, each is distinguished from the other by a number of characteristic innovations. Phonological innovations on the Indian side include the loss of the Indo-Iranian diphthongs *\*ai*, *\*au* (> *e*, *o*) and voiced sibilants (*\*z*, *\*ž*, etc.) and the development of a series of retroflex consonants (*t*, *ṇ*, *ṣ*, etc.), whilst Iranian

languages typically show the loss of the voiced aspirates \*b<sup>h</sup>, \*d<sup>h</sup>, \*g<sup>h</sup>, etc. (> *b*, *d*, *g*), the development of the voiceless fricatives *f*, *θ*, *x* (from IIr. \*p, \*t, \*k before consonants, and from IIr. \*p<sup>h</sup>, \*t<sup>h</sup>, \*k<sup>h</sup>), the depalatalization of IIr. \*ć, \*j<sup>(h)</sup> (> \*ʃ, \*dʒ, OPers. *θ*, *d*, Av. *s*, *z*) and the change of \*s (in most positions) to *h*. Some apparent exceptions to these isoglosses may be due to the reversal of a sound-change: for instance, Av. *pt* (as in *hapta* ‘seven’) may derive from the expected \*ft (as attested, directly or indirectly, in all other Iranian languages, e.g. Pers. *hafi*) rather than preserving IIr. \*pt (cf. OInd. *saptá*). In other cases, however, it is clear that a development characteristic of Iranian cannot in fact have been fully carried through at the Common Iranian stage: cf., for instance, p. 271 below on evidence for the survival of the palatal \*ć in certain clusters. Similarly, the development of \*s to *h*, though common to Avestan, Old Persian and all later Iranian languages, is demonstrably later than the earliest attestations of Iranian in Ancient Near Eastern sources (cf. the next paragraph on the divine name *Assara mazaš*). Thus, at least in phonology, the innovations attributable to Common Iranian are comparatively few in number (though significant in kind).

The original homeland of the Aryans, the speakers of Common Indo-Iranian, cannot be precisely identified, but is thought to have been in western Central Asia, to the east and northeast of the Caspian Sea. At a time when “Proto-Indian” and “Proto-Iranian” (i.e. the ancestral dialects from which the Indian and Iranian languages respectively derive) had already become differentiated to some extent, perhaps about the beginning of the 2nd millennium B.C.E., two groups of “Proto-Indian” or “Indo-Aryan” speakers began to migrate from this homeland, one towards the west (cf. above, p. 205, on traces of the Indo-Aryans in the Hurrian empire of Mitanni in northern Mesopotamia) and the other towards India. At a later date, Iranian tribes too began to migrate westwards, reaching central and western Iran by the middle of the ninth century B.C.E., at which period they are referred to for the first time in Assyrian sources; whether they had come from the northeast by the most direct route, to the south of the Caspian, or more circuitously through the Caucasus is still uncertain. (For a more detailed summary and references to the literature on the prehistory of the Aryans see Schmitt 1987.) From the ninth century B.C.E. onwards, a scattering of Iranian linguistic material is to be found in Mesopotamian sources, beginning with the names of the Medes (*Matai*) and Persians (*Parsuaš*) and most notably including the name of the principal deity of the Iranians in the form *Assara mazaš* (= Common Iranian \*Asura-mazdās, later \*Ahura-mazdāh, cf. OPers. *Auramazdā*, Av. *Ahurō Mazdā*).

It is in part as a result of the fact that Common Iranian cannot have differed greatly from Common Indo-Iranian (or “Aryan”) that it is difficult to determine the precise status of the so-called Nuristani languages (formerly known as “Kafiri”). This group of languages, recorded in modern times in the northeast of Afghanistan and neighbouring parts of Pakistan, undoubtedly belongs to the Indo-Iranian family, but the nature of the relationship is not clear. All three theoretically possible solutions have been advocated: a third, independent sub-group beside Indian and Iranian (Morgenstierne 1973: 327–343); an archaic form of Iranian, much influenced by several millennia of proximity to languages of the Indian group (Mayrhofer 1983); or a branch of “Proto-Indian” which split off before the arrival of the Indo-Aryans in India and subsequently developed in intensive contact with Iranian languages (Degener 2002).

Only two Old Iranian languages are attested by texts, namely Avestan and Old Persian. Others, such as Median and Scythian, are known to us only through occasional words and names transmitted in texts in other languages.

Avestan is the language of the Zoroastrian scriptures, the Avesta, the earliest parts of which are the Gāthās (“Songs”) of Zoroaster or Zarathushtra – whom tradition places in the sixth century B.C.E., though many scholars argue, partly on linguistic grounds, for a date five centuries or more earlier – and the *Yasna Haptanḫāti* “Service consisting of seven chapters”. These texts,

together with a few short prayers, are preserved in Old or “Gathic” Avestan, a highly archaic dialect comparable to Vedic in its stage of development. Later Avestan, also known as Younger Avestan, is attested by a much larger corpus of texts, including the Yashts (hymns in honour of individual divinities) and the *Vidēvdād* (“Law against the demons”). The manuscript tradition of the Avesta goes back to the Sasanian period (224–651 C.E.), when these orally transmitted texts were written down, probably for the first time, in a specially invented and extremely elaborate alphabetic script (see Hoffmann & Narten 1989 and the script table below).

Schritttafel der Avesta-Schrift

1 𐬀	2 𐬁	3 𐬂	4 𐬃	5 𐬄	6 𐬅 𐬆	7 𐬇	8 𐬈		
<i>a</i>	<i>ā</i>	<i>â</i>	<i>ã</i>	<i>ȧ</i>	<i>â̇</i>	<i>ə</i>	<i>ē</i>		
		9 𐬉	10 𐬊	11 𐬋	12 𐬌	13 𐬍	14 𐬎	15 𐬏	16 𐬐
		<i>e</i>	<i>ē</i>	<i>o</i>	<i>ō</i>	<i>i</i>	<i>î</i>	<i>u</i>	<i>û</i>
17 𐬑	18 𐬒	19 𐬓	20 𐬔	21 𐬕	22 𐬖	23 𐬗			
<i>k</i>	<i>x</i>	<i>â</i>	<i>x<sup>v</sup></i>	<i>g</i>	<i>ġ</i>	<i>γ</i>			
24 𐬘	25 𐬙								
<i>c</i>	<i>j</i>								
26 𐬚	27 𐬛	28 𐬜	29 𐬝 𐬞	30 𐬟 𐬠					
<i>t</i>	<i>θ</i>	<i>d</i>	<i>ð</i>	<i>ṫ</i>					
31 𐬡	32 𐬢	33 𐬣	34 𐬤						
<i>p</i>	<i>f</i>	<i>b</i>	<i>β</i>						
35 𐬥	36 𐬦	37 𐬧	38 𐬨	39 𐬩 𐬪	40 𐬫	41 𐬬	42 𐬭		
<i>ŋ</i>	<i>ñ</i>	<i>ŋ<sup>v</sup></i>	<i>n</i>	<i>ñ</i>	<i>ṅ</i>	<i>m</i>	<i>ṁ</i>		
43 𐬮	44 𐬯	45 𐬰 𐬱	46 𐬲						
<i>y</i>	<i>γ</i>	<i>v</i>	<i>r</i>						
47 𐬳	48 𐬴	49 𐬵	50 𐬶	51 𐬷	52 𐬸				
<i>s</i>	<i>z</i>	<i>š</i>	<i>ž</i>	<i>ṧ</i>	<i>ž̇</i>				

53  
𐬹 𐬺  
*h*

The Avestan orthography was designed to preserve the traditional pronunciation with great accuracy and contains much phonetic detail which is irrelevant for the comparativist. For example, the word for 'land' appears in such different forms as *da'hju-* and *da'xiu-*, both representing what is etymologically and probably phonemically \*dahyu-. A particularly confusing feature of the Avestan writing system is the frequent notation of anaptyctic and epenthetic vowels. In the present chapter such unetymological vowels will be written in superscript, as in *daδā<sup>h</sup>ti* 'he gives' (= OInd. *dādāti*) – contrast the diphthong *āi* in *āiḍi* 'come!' – or GAv. *d<sup>h</sup>bitiia-* 'second' (= OInd. *dvi<sup>h</sup>ya-*). Note too that the semi-vowels *y* and *w* are regularly represented by *ii* and *uu* (which can equally represent the sequences *iy* and *uw*) and that *ī*, *ū* are not consistently distinguished from *i*, *u*.

Old Persian, which is known from inscriptions of the Achaemenian period (sixth to fourth centuries B.C.E.), represents a later stage of linguistic development as well as a different dialect from the language of the Avesta. Like Avestan, it is written in a specially invented script, in this case a form of the cuneiform writing commonly used in the Ancient Near East. The Old Persian script combines syllabic and alphabetic principles. For example, there are two *t*-signs, of which *t<sup>h</sup>* is syllabic (representing [tu(:)], since *ī* and *ū* are not distinguished from *i* and *u*) whilst *t* can represent either a syllable [ta, tə] or the simple consonant [t]. Since there is no sign for \*t<sup>i</sup> (though comparable signs such as *d<sup>h</sup>* do exist), [ti] or [ti:] has to be written by means of two signs (*t-i*), a combination which can also denote [tai]. The fact that a sign such as *t* has both syllabic and consonantal values is the source of much ambiguity, as is the fact that in most cases a nasal is not written before another consonant. As a result of these two deficiencies, for instance, the 3 sg. pres. ind. act. and mid. (*-ti* and *-tai*) and the equivalent pl. endings (*-nti* and *\*-ntai*) are all indistinguishable in writing. In this chapter, for the sake of clarity, Old Persian forms will generally be cited in phonemic transcription rather than in transliteration. (On the Old Persian writing system see further Hoffmann 1976: 620–645.)

Only a brief survey can be given here of the great variety of languages attested at the Middle Iranian stage. Western Middle Iranian is represented by Middle Persian, which is essentially, though not in every detail, a later form of the same dialect as Old Persian, and by Parthian. The Eastern Middle Iranian languages include Khotanese and the closely related Tumshuqese, which are the most conservative of the Middle Iranian languages in their morphology, Sogdian, Bactrian and Choresmian. Amongst the even more numerous Modern Iranian languages we shall occasionally have reason to refer to Persian (or New Persian), Pashto, Ossetic and the Shughni group. Further information on these and other Iranian languages may be found in the relevant chapters of the *Compendium Linguarum Iranicarum* (Schmitt 1989); on the Middle Iranian languages see also Henning 1958.

## PHONOLOGY

The vocalic system of Common Iranian is almost identical to that of Old Indian. The main difference is the lack of *ē* and *ō*, the diphthongs from which OInd. *e* and *o* derive being preserved in Old Iranian as *ai* (Av. *aē* or *ōi*) and *au* (Av. *ao* or *ōu*). Cf. OPers. *daiva-*, Av. *daēuua-* '(evil) god, devil' = OInd. *devá-* 'god'; OPers. *rautah-* 'river' = OInd. *sró-tas-* 'stream'. The comparatively rare long diphthongs *āi* and *āu*, which are shortened in Indian, also survive in Old Iranian, cf. Av. instr. pl. *yasnāiš* 'by sacrifices' = OInd. *yajñāiš*; Av. nom. sg. *gāuš* 'ox, cow' = OInd. *gāuṣ*.

The etymological origins of the Iranian vowels *a*, *i*, *u*, *ɣ* (= [ɤ<sup>h</sup>]), *ā*, *ī*, *ū* are in general the same as those of the equivalent Old Indian vowels. In particular, as in Indian, PIE \*a, \*e, \*o, \*ṛ, \*ṝ fall together as *a*, and the corresponding long vowels (including those

which ultimately derive from short vowel + laryngeal) as  $\bar{a}$ . Brugmann's Law, according to which PIE \*o gives  $\bar{a}$  in open syllables, seems to apply in the same circumstances in both branches of Indo-Iranian, e.g. Av. nom./acc. sg. *dāru*, OInd. *dāru* 'wood' = Gr. δόρυ. However, the contexts in which the PIE laryngeals are vocalized (to Iran. *i*, OInd.  $\bar{i}$ , e.g. OPers./Av. *pitar-*, OInd. *pitar-* 'father' < PIE \*ph<sub>2</sub>ter-) are more restricted in Iranian, resulting in many cases of the correspondence Iran. 0 : OInd.  $\bar{i}$ , e.g. GAv. *dug<sup>h</sup>dar-*, LAv. *duγ<sup>h</sup>dar-* = OInd. *duhitār-* 'daughter' < PIE \*d<sup>h</sup>ugh<sub>2</sub>ter-; GAv. *vər<sup>h</sup>ntē* = OInd. *vṛ<sup>h</sup>ntē* 'chooses' < \*w<sub>1</sub>-nH-toy. Where Old Indian has *ir*, *ur* (before vowels) or  $\bar{i}r$ ,  $\bar{u}r$  (before consonants) from PIE \*r̥H or \*r̥H, Iranian has uniformly *ar*:

- Av. *sarah-* = OInd. *śiras-* 'head' < \*k<sup>h</sup>ṛ<sub>2</sub>os-  
 OPers. *paru-* = OInd. *purú-* 'much' < \*p<sup>h</sup>h<sub>1</sub>u-  
 OPers. *darga-*, Av. *dar<sup>h</sup>ga-* = OInd. *dārg<sup>h</sup>á-* 'long' < \*d<sup>h</sup>h<sub>1</sub>g<sup>h</sup>o-  
 Av. *var<sup>h</sup>nā-* = OInd. *ūr<sup>h</sup>nā-* 'wool' < \*h<sub>2</sub>w<sup>h</sup>h<sub>1</sub>neh<sub>2</sub>-

Since short *a* and long  $\bar{a}$  probably differed markedly in quality as well as in quantity (as they do both in Sanskrit and in many Modern Iranian languages), the system of simple vowels in Common Iranian may be represented diagrammatically as in Figure 4.3.

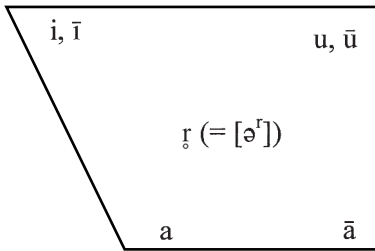


FIGURE 4.3 SIMPLE VOWELS

This simple system seems to survive almost unchanged in Old Persian (in so far as the inadequate cuneiform orthography allows one to tell), though the unitary sound *r̥* had probably developed into a sequence of vowel + consonant, most likely [ər] (written *a-r-* in initial position, but distinct from the sequence [ar], as is proved by its different fate in Middle and New Persian). A similar development is found in Avestan, where *r̥* usually gives *əṛə* (i.e. *əṛ<sup>h</sup>*) as in *kəṛ<sup>h</sup>nao<sup>h</sup>ti* 'makes' = OInd. *kṛ<sup>h</sup>ṇóti*. But many other outcomes of *r̥* are found in Avestan, e.g. *ar* before *š* (*arš<sup>h</sup>ti-* 'spear' = OPers. *a-r-š-t-i-* [ərš<sup>h</sup>ti-], OInd. *ṛ<sup>h</sup>ṣ<sup>h</sup>tī-*), *əhr* before *k*, *p* (*vəhrka-* 'wolf' = Pers. *gurg* < OPers. \*vərka-, OInd. *vṛ<sup>h</sup>ka-*, cf. below, p. 272–273), *ir* before *y* (pres. stem *kiriia-* 'to be done' = OPers. *kəriya-* < \*kṛya-, OInd. *kriyá-*), *rə* after *t* (pres. stem *trəfiia-* 'to steal' = Sogd. *čəf-* < \*tṛfya-, cf. OInd. *ṛ<sup>h</sup>f<sup>h</sup>p-* 'stealing').

The example of *r̥* may serve to illustrate the complexity of Avestan vocalism. Considerations of space make it impossible to follow in equal detail the development of all the vowels (for which see de Vaan 2003), but the most important contextual changes must be noted. These include the (re)appearance of mid-high vowels  $\bar{e}$  and  $\bar{o}$ . In final position Ilr. \*-ai and \*-au give  $\bar{e}$  and  $\bar{o}$  (or *-uuō*) respectively (e.g. 3 sg. pres. mid. ending *-tē* = OPers. *-tai*, OInd. *-te* < \*-tai, PIE \*-toy), while \*-yā gives LAv. *-e* (e.g. *a*-stem gen. sg. ending LAv. *-ahe* = GAv. *-ahiiā*, OPers. *-ahəyā*, OInd. *-asya*, PIE \*-osyo). Internally, *a* often becomes *e* between two palatal sounds, as in GAv. *yehiiā*, LAv. *yejhe* 'of whom'



(= OInd. *yásya*), and *o* between *p/g/m/v* and *u*, as in *po<sup>h</sup>ru-* ‘much’ (= OPers. *paru-*, cf. above on this page). Final *-ō* (GAv. *-ō* or *-ā*) and *-ā* most often derive, via *\*-ah* and *\*-āh*, from *\*-as* (PIE *\*-os*, *\*-es*) and *\*-ās* respectively, e.g. *yō* (GAv. *yā*) ‘who’ (nom. sg. m.), *yā* ‘who’ (nom. pl. f.). Before nasals, especially in final syllables, *a* and *ā* normally develop into *ə* and *q* (= [ã]) respectively, so that *-əm* and *-qm* are the regular acc. sg. endings of *a*-stems and *ā*-stems. The *ə* which arises thus is subject to further changes, for instance to *i* after a palatal, as in *°činah-* ‘desire’ < *\*čənah-* (beside OPers./Av. *°čanah-*, OInd. *cānas-*). The sequences *\*(i)yə*, *\*(u)wə* frequently contract to *ī*, *ū* (or *i*, *u*, since the length of these vowels is not consistently noted in Avestan), *\*ayə*, *\*awə* to *aē*, *ao*. Cf. *īm* ‘this’ (= OPers. *iyam*, OInd. *iyám*, nom. sg. f.); LAv. *tūm* ‘thou’ (= GAv. *tuiōm*, OPers. *tuvam*, OInd. *tuvám*, nom.); *aēm* ‘this’ (beside GAv. *aiiōm*, OInd. *ayám*, nom. sg. m.); *baon* ‘they became’ (= OInd. (*ā*)*bhavan*).

The consonantism of the Iranian languages diverges much more fundamentally from that of Old Indian. Two major innovations in Iranian are the loss of all aspirates and the appearance of a series of fricatives (*f*, *θ*, *x*) unknown to Old Indian. In most cases these fricatives derive from *p*, *t*, *k* in preconsonantal position, but they also correspond to the Old Indian voiceless aspirates *ph*, *th*, *kh* (in all positions). The voiced aspirates (PIE *\*b<sup>h</sup>*, *\*d<sup>h</sup>*, *\*g<sup>h</sup>/g<sup>wh</sup>*, OInd. *bh*, *dh*, *gh*) merely lose their aspiration, thus merging with the original non-aspirate series:

Iran. <i>p</i>	= OInd. <i>p</i> : OPers./Av. <i>pitar-</i> , OInd. <i>pitár-</i> ‘father’
Iran. <i>f</i>	= OInd. <i>p</i> : Av. <i>friia-</i> , OInd. <i>priyá-</i> ‘dear’ = OInd. <i>ph</i> : Av. <i>kafa-</i> ‘foam’, OInd. <i>kapha-</i> ‘slime’
Iran. <i>t</i>	= OInd. <i>t</i> : OPers. <i>tuvam</i> , GAv. <i>tuiōm</i> , LAv. <i>tūm</i> , OInd. <i>tuvám</i> (nom.) ‘thou’
Iran. <i>θ</i>	= OInd. <i>t</i> : OPers. <i>θuvām</i> , Av. <i>θβqm</i> , OInd. <i>tvám</i> (acc.) ‘thee’ = OInd. <i>th</i> : Av. <i>paθō</i> , OInd. <i>pathás</i> (gen. sg.) ‘way’
Iran. <i>k</i>	= OInd. <i>k</i> : Av. <i>kuθra</i> , OInd. <i>kútra</i> ‘whither?’
Iran. <i>x</i>	= OInd. <i>k</i> : OPers./Av. <i>xšap-</i> , OInd. <i>kšap-</i> ‘night’ = OInd. <i>kh</i> : Av. <i>xā-</i> , OInd. <i>khā-</i> ‘spring, well’
Iran. <i>b</i>	= OInd. <i>b</i> : Ossetic <i>bal</i> ‘group’, OInd. <i>bála-</i> ‘power’ (?) = OInd. <i>bh</i> : OPers./Av. <i>brātar-</i> , OInd. <i>bhrātár-</i> ‘brother’
Iran. <i>d</i>	= OInd. <i>d</i> : Av. <i>dāntan-</i> , OInd. <i>dant-</i> ‘tooth’ = OInd. <i>dh</i> : Av. <i>daēnu-</i> ‘female’, OInd. <i>dhenú-</i> ‘cow’
Iran. <i>g</i>	= OInd. <i>g</i> : Av. <i>ga’ri-</i> , OInd. <i>girí-</i> ‘mountain’ = OInd. <i>gh</i> : Av. <i>gar<sup>h</sup>ma-</i> , OInd. <i>gharmá-</i> ‘heat’

Despite their loss of aspiration in Iranian, the PIE voiced aspirates are still occasionally distinguishable from the equivalent non-aspirates by the effects of Bartholomae’s Law (cf. above, p. 229), according to which a combination such as *\*g<sup>h</sup>+t* was assimilated to *\*gd<sup>h</sup>* in Indo-Iranian (and perhaps already in Indo-European), whereas *\*g+t* gave *\*kt*. By this rule, which applied to all combinations of voiced aspirate + voiceless stop or sibilant, one may deduce from a form such as GAv. *aog<sup>h</sup>dā* ‘he said’ (= *aog* + morpheme *-tā*) that the root *aog* originally ended in an aspirate *\*g<sup>h</sup>* or *\*g<sup>wh</sup>*; in this case *\*g<sup>h</sup>* is confirmed by Gr. εὔχασθαι, etc. Unfortunately, the contrary deduction cannot usually be made from the presence of a voiceless cluster, since the effects of Bartholomae’s Law tended to be cancelled out by the restoration of the normal form of the morpheme, as in LAv. *aoxta* for GAv. *aog<sup>h</sup>dā* or OPers./LAv. *basta-* ‘bound’ for expected *\*bazda-* (= OInd. *baddhá-*).

Common Iranian was rich in both sibilants (*s*, *z*, *š*, *ž*) and affricates (*č*, *č̣*, i.e. [tʃ, dʒ]) – differing from the OInd. *c*, *j*, which, at least in the earliest period, were palatal stops – and

possibly \*c, \*j, i.e. [ʦ, ʤ]). These stem in part from PIE \*s and in part from the “two series of palatals”, i.e. (i) PIE \*k, \*ǵ, \*ǵʰ (> OInd. ś, j, h) and (ii) PIE \*k/kʷ, \*g/gʷ, \*gʰ/gʷʰ when secondarily palatalized before PIE \*e, \*i, \*y, etc. (> OInd. c, j, h). The history of these sounds is rather complicated, but is worth examining in some detail in view of the fact that Iranian here retains evidence of distinctions which are lost in Old Indian.

For Indo-European only one sibilant is to be assumed, namely \*s (with the allophone \*z in clusters such as \*zd). In addition to its role as an independent phoneme, PIE \*s has a secondary origin as an automatic feature of the juncture of two dental stops (cf. above, p. 54): \*t+t/\*d+t = \*tst, \*dʰ+t = \*dzdʰ, e.g. PIE \*sed+to- = \*setsto- ‘seated’ > OInd. *sattá-*, MPers. [*ni*]šast (< OPers. \*[ni]šasta-), Lat. *sessus*, cf. also OIr. *sess* ‘seat’, etc.; \*wǵdʰ+to- = \*wǵdzdʰo- ‘increased’ > OInd. *vṛddhá-*, Av. *vər̥zda-* (the development of a voiced group in the latter case being a further instance of Bartholomae’s Law). As these examples indicate, the resulting clusters were simplified in different ways in Indian, where the sibilant disappeared, and in Iranian, where the first of the two dental stops was lost, giving the regular correspondences OInd. *tt* : Iran. *st* and OInd. *ddh* : Iran. *zd*.

In Indo-Iranian, much as in Slavic (cf. below, p. 526–527), PIE \*s and \*z underwent a split, becoming Indian ś, \*z, Iran. š, ž after the sounds collectively known as “RUKI” (i.e. *r, ʀ; ũ, āu; k* and other velars and palatals; *ř, āi*) but remaining, at least in the first instance, unchanged in other contexts. Examples: the loc. pl. ending Av. *-šu* (OInd. *-ṣu*) after stems in *u* (etc.) but *-su* after stems in *ant*; Av. *mižda-* ‘reward’, OInd. *mīdhá-* < \*mizdʰá-, PIE \*miz-dʰ(h)yo- (Gr. μισθός) but OPers./Av. *Mazdā-* (divine name) < PIE \*mǵz-dʰeh₁-, cf. OInd. *medhá-* ‘wisdom’ < PIE \*mǵz-dʰeh₁-. (Note that the Iranian forms with *z* and *ž* here clarify their Indian counterparts, which have become opaque as a result of the loss of voiced sibilants in Old Indian.) This change does not affect Iran. *st, zd* < \*tst, \*dzdʰ, showing that the sibilant was still protected by the preceding stop when the RUKI rule operated: Av. *vista-* ‘known’ < \*witsto- = \*wid+to- (Gr. φισθός, OIr. *fess*). In Iranian (but not Indian) the change to š, ž also takes place after a labial: Av. *dižza-*, OInd. *dīpsa-* < \*di(d)bzʰa-, desiderative of Av. *dab*, OInd. *dabh* ‘to injure, deceive’.

Finally, those instances of PIE \*s which had so far survived unchanged in Iranian underwent a further split, *s* remaining in groups such as *sn, sp, st, \*ts* (> *s*) but becoming *h* in all other contexts, e.g. OPers. *a(h)mi*, Av. *ahmi*, Khot. *imā* ‘I am’, OPers. *hanti*, Av. *hənti*, Khot. *indā* ‘they are’ (= OInd. *ásmi, sánti* < PIE \*h₁esmi, \*h₁senti), but OPers. *asti*, Av. *asti*, Khot. *ástā* ‘he is’ (= OInd. *ásti* < PIE \*h₁esti); *ā-*stem loc. pl. Av. *-ā-hu* (= OInd. *-ā-su*). Although this development is found in all Iranian languages it must be comparatively late, since Proto-Iranian forms with *s* (for later *h*) are preserved in ancient Near Eastern sources (cf. above, p. 264).

An important implication of this fact is that the development of the PIE “first palatal series” (\*k, \*ǵ, \*ǵʰ) to sibilants (*s, z*), which occurs in Avestan and all branches of Iranian other than Old Persian (and later dialects of southwestern Iran), must also be later than the Common Iranian period, since the *s* arising from PIE \*k does not participate in the change of PIE \*s to *h*. As a plausible intermediate stage between the attested Iranian series (Av. *s, z, z*, OPers. *θ, d, d*) and the presumed Indo-Iranian palatal affricates \*č, \*j, \*jʰ (< PIE \*k, \*ǵ, \*ǵʰ) the dental affricates \*ʦ and \*ʤ may be reconstructed for Common Iranian. In Figure 4.4 the postulated development of this “first series” of palatals is set beside that of the “second series” (i.e. the Indo-Iranian palatal stops arising from the secondary palatalization of PIE velars or labio-velars, which eventually gave palatal affricates in Iranian) in order to show how the resulting Iranian and Indian forms disambiguate one another.

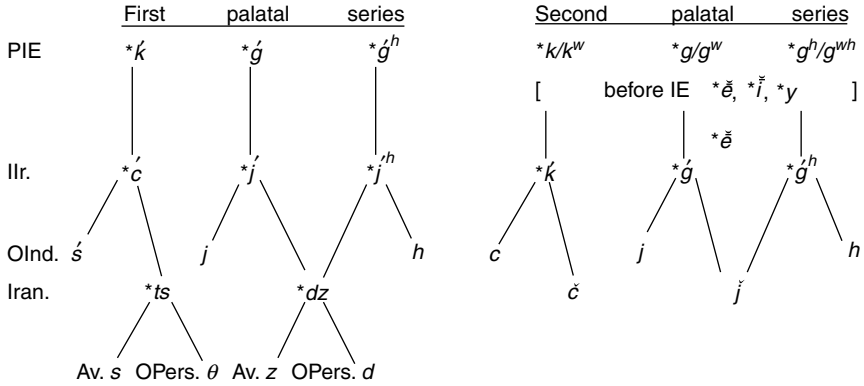


FIGURE 4.4 THE FIRST AND SECOND SERIES OF PALATALS IN IRANIAN

As Figure 4.4 shows, only the voiceless sounds (OInd. ś, Iran. \*ts < PIE \*k; OInd. c, Iran. č < PIE \*k/k<sup>w</sup> before a palatal) are etymologically unambiguous. Each of the voiced sounds, OInd. j and h, Iran. \*dz and j̄, has a double origin, since Old Indian confuses the two palatal series while Iranian (as always) confuses aspirates and non-aspirates. However, the ambiguity is resolved in the case of words preserved in both branches of Indo-Iranian, each of which preserves the distinction lost in the other:

- Iran. \*ts = OInd. ś: OPers. θard-, Av. sar<sup>o</sup>d-, cf. OInd. śarād- ‘year’ (PIE \*k)
- Iran. \*dz = OInd. j: OPers. yad, Av. yaz, OInd. yaj ‘to worship’ (PIE \*ǵ)
- = OInd. h: OPers. daraniya-, Av. zaraniia-, OInd. hiraṇya- ‘gold’ (PIE \*ǵ<sup>h</sup>)
- Iran. č = OInd. c: OPers. či, Av. čit̄, OInd. cit (enclitic) ‘also, even’ (PIE \*k<sup>w</sup>)
- Iran. j̄ = OInd. j: Av. jā’ni-, OInd. jāni- ‘woman’ (PIE \*g<sup>w</sup>)
- = OInd. h: OPers./Av. jan, OInd. han ‘to strike, kill’ (PIE \*g<sup>wh</sup>)

The depalatalization seen in the unconditioned reflexes of the PIE palatals (Av. s, z, OPers. θ, d) failed to take place in most combinations with consonants, where the usual outcome in all Iranian languages is palatal ś, ž, as in Av. fšū- beside pasu- ‘sheep’ (OInd. paśú-, PIE \*p(e)ku-). In most cases the Old Indian equivalent is retroflex ś, \*z̄, cf. Av. ašta ‘eight’ = OInd. aṣṭā (PIE \*Hoktoh<sub>1</sub>); Av. mərəždika- ‘mercy’ = OInd. mṛḍikā- (< \*mṛzḍikā-, PIE \*ǵd-). An important special case is that of PIE \*sk, which gives OInd. (c)ch, Iran. s, as in the inchoative present stem OInd. gáccha- ‘to come’, Choresmian [n]ʃs- < \*[ni]gasa- ‘to arrive’ (a more archaic form than Av. jasa-), all < PIE \*g<sup>w</sup>m̄-sko- (Gr. βᾶσκε). Regarding PIE \*kw (> OPers. s, Khot. śś [ʃ], elsewhere sp), \*ǵ<sup>h</sup>(w) (> OPers. z, Khot. ś [ʒ], elsewhere zb) and \*kr see below, p. 271.

Finally, we should note the outcome of PIE clusters of velar, labio-velar or palatal + \*s. All such groups give OInd. kṣ, whilst Iranian distinguishes four possibilities: (i) xš < \*k<sup>(w)</sup>s, e.g. Av. vaxšiiia-, OInd. vakṣyá-, future of vak ‘to speak’; (ii) gž < \*g<sup>(w)</sup>z<sup>h</sup> (for \*g<sup>(w)</sup>h+s by Bartholomae’s Law), e.g. GAV. aog<sup>o</sup>žā ‘saidst’; (iii) š < \*ks, e.g. Av. mošu, OInd. makṣú ‘quickly’; and (iv) ž < \*ǵz<sup>h</sup> (for \*ǵ<sup>h</sup>+s), e.g. GAV. dīdā<sup>o</sup>žā-, desiderative of dar<sup>o</sup>z ‘to make firm’ (OInd. dṛ̥mhati).

The following schema shows the minimum complement of consonantal phonemes to be assumed for Common Iranian. An asterisk (\*) indicates those which do not survive as such in any attested language.

p	f	b		m	w
t	θ	d		n	l
	*ts	s	*dʒ	z	r
	č	š	ǰ	ž	y
k	x	g			*H

Regarding the reappearance of PIE \*l (> OPers./Av. *r*) as *l* in later Iranian see the next paragraph. On the reconstructions \*ts and \*dʒ see above, p. 269–270. The symbol \*H represents a consonant deriving from the PIE laryngeals, whose survival, at least in certain positions, is indicated by metrical and other considerations; e.g. GAv. *mazdā*, a form which is disyllabic as the nom. sg. but trisyllabic as the gen. sg., indicating nom. \*mazdaH-s (> \*mazdās) ~ gen. \*mazdaH-as.

Not all of the phonological developments shared by Avestan and Old Persian can be ascribed to Common Iranian. The change of \*s to *h* (except in certain groups), which occurs in all attested Iranian languages, cannot have been completed until after the arrival of Iranian speakers in western Iran, as has already been pointed out. The replacement of \*l (and \*l̥) by *r* (and *r̥*), which Avestan and Old Persian have in common with Vedic, was nevertheless not universal in Iranian, as is shown by the later reappearance of dialectal forms with *l* < PIE \*l, e.g. Pers. *lištan* ‘to lick’ beside Av. *raēz* (PIE \*leyǵ<sup>h</sup>-, Gr. λείγω). Similarly, the two Old Iranian languages share a development of PIE \*k<sup>(w)</sup>y to *šy*, as in OPers. *šiyav*, Av. *š(ii)auu* ‘to go’ (PIE \*kyew-, OInd. *cyav-*, Gr. σέδομαι and κινέω), but the preservation of an affricate in Khot. *tsū-* [ts<sup>h</sup>u-], Tumshuqese *cch-* ‘id.’ indicates that only the intermediate stage \*čy is to be attributed to Common Iranian.

The most important isogloss separating Old Persian from Avestan is to be seen in the treatment of the “first palatal series”, PIE \*k, \*ǵ, \*ǵ<sup>h</sup>, which are thought to have developed via palatal affricates (Iir. \*č, \*j, \*j<sup>h</sup>) and dental affricates (Common Iranian \*ts, \*dʒ) and which give *θ* and *d* in Old Persian (and later dialects of southwestern Iran) but *s* and *z* in Avestan and all other Iranian languages (cf. above, p. 269 ff.). The treatment of the PIE combinations \*kw and \*ǵ<sup>(h)</sup>w provides a three-way isogloss, giving *sp*, *zb* in most Iranian languages (including Avestan), *s*, *z* in Old Persian, and *š*, *ž* in the group of northeastern Iranian Saka (Scythian) languages represented by Khotanese. Examples: Av. *aspa-*, OPers. *asa-*, Khot. *aśśa-* [aʃa-] ‘horse’ (= OInd. *ásva-*, PIE \*-kw-); Parthian *əzbān*, OPers. *həzan-*, Khot. *biśāa-* [βiʒāa-] ‘tongue’ (cf. OInd. *jihvā-*, PIE \*-ǵ<sup>h</sup>w-). Since the palatals *š*, *ž* can hardly be derived from \*tsw and \*dzw, it is simplest to assume Common Iranian \*čw and \*jw. The palatal nature of Iir. \*č < PIE \*k seems also to have been preserved up to the Common Iranian stage in the case of the cluster \*čr, cf. Khot. *śśāra-* [ʃəra-] ‘good’ (= Av. *srīra-*, OInd. *śrīla-* ‘beautiful’, cf. Gr. κρείων). In Old Persian \*čr gives *ç* (a sibilant of unclear phonetic character), a development which may have proceeded via \*tsr and \*θr, since *ç* is also the outcome of Iran. \*θr < PIE \*tr or \*tl, as in *puça-* ‘son’ (= Av. *puθra-*, OInd. *putrá-*).

It is not surprising to find that it is the languages spoken at the fringes of the Iranian world – Old Persian in the extreme southwest and the languages of the nomadic Saka peoples of the Eurasian steppes – which stand out as aberrant in respect of the old isoglosses mentioned above. In each case, Avestan represents the Iranian mainstream. Avestan is often regarded as an Eastern Iranian language, which is no doubt correct from a purely geographical point of view, but it shows none of the phonological developments which are characteristic of Eastern Iranian in later periods, such as the voicing of the fricative in the groups \*xt and \*ft or the depalatalization of \*č. Avestan does indeed have its peculiarities, such as the reversion of \*ft to *pt* (cf. above, p. 264), the development of \*rt

to *š* (cf. below, p. 272–273) or the frequent insertion of a nasal *ŋ* before *h* (e.g. *ajhaṭ*, 3 sg. subj. of *ah-* ‘to be’, OInd. *ásat*), but they do not seem likely to be very ancient, nor do they provide evidence of a particularly close relationship with any other Iranian language.

## MORPHOPHONOLOGY

At the end of the word certain special phonological changes take place. In the attested Old Iranian languages the original distinctions between long and short final vowels are lost. In general, Old Persian and Old Avestan tend to lengthen short final vowels, while Later Avestan shortens many that were originally long. In the *a*-declension, for instance, both the voc. sg. (originally *\*-a*) and the instr. sg. (originally *\*-ā*) appear as OPers./GAv. *-ā*, LAv. *-a*, so that the length of the final vowel no longer has any phonemic (or etymological) significance. The merging of long and short final vowels was not universal, however; cf. Morgenstierne 1973: 108–9 on remnants of a distinction between *\*-a* and *\*-ā* in Shughni and other Modern Iranian languages of the Pamir mountains.

A feature common to all the Iranian languages is the loss of final *\*-h* (< PIE *\*-s*). In some languages the loss of *\*-h* is accompanied by a change in the quality of the preceding vowel, whereby *\*-ah* > Av. *-ō* (GAv. also *-ə*), Khot. *-ā* [-e], Sogd. *-i*, and *\*-āh* > Av. *-ā*, Khot. *-e* [-e:] (but Sogd. *-a*; cf. the similar changes accompanying the loss of final *\*-m* in Middle Iranian: *\*-am* > Khot./Sogd. *-u*; *\*-ām* > Khot. *-o* but Sogd. *-a*). In Old Persian, on the other hand, *\*-h* is lost without a trace, as are *\*-d/-t* and perhaps some other final consonants, so that *\*-ah/-ad* and *\*-āh/-ād* give *-a* and *-ā* respectively (thus re-establishing the recently lost phonemic distinction between long and short final vowels). Such developments had a significant impact on the morphology of the Iranian languages, as may be seen from the paradigm of the *a*-stems in Table 4.21.

The changes typical of absolute word-final position are sometimes found also internally, in compounds and before particular morphemes: cf. Av. *vačō.mar<sup>ta</sup>* ‘recited aloud’ and instr. pl. *vačēbiš*, both from *vačah-* ‘speech, word’ with the same treatment of *\*-ah* < *\*-as* as occurs in final position in the nom./acc. sg. *vačō/vačē* (= OInd. *vācas*, Gr. (F)ῥπος). In other cases, however, compound-juncture is treated as internal position, as in Av. *vačas.tašti-* ‘strophe’, where the original *\*s* “reappears” in accordance with the regular treatment of the PIE cluster *\*st*. Such combinatory variants as *vačas<sup>o</sup>* are referred to as sandhi-forms, “sandhi” being the Sanskrit term for the “combination” both of elements within a word and of words within a sentence (cf. above, p. 229–230); in Old Iranian, however, the occurrence of sandhi is almost entirely restricted to the juncture of elements within a single accentual unit, i.e. of morphemes in a word, of words in a compound or of a clitic with its host as in Av. *fradātaē-ča* ‘and (it) will prosper’ (= *\*frāda<sup>te</sup>*, 3 sg. subj. mid. of *frād* + encl. *-ča* ‘and’, cf. Hoffmann 1975: 262ff.), *kas-čit* ‘someone’ (= nom. sg. m. *kō* ‘who?’ + encl. indefinite ptcl. *-čit*), OPers. *kaš-či*. As these examples show, the forms occurring in sandhi before enclitics often preserve older phonological forms of the inflections: *-taē<sup>o</sup>* < PIE *\*-toy* (cf. above, p. 267), *kas<sup>o</sup>* < PIE *\*k<sup>w</sup>os*. The shortening of the vowel in the first syllable of *fradātaē-ča* is probably due to a shift of accent to the syllable preceding the enclitic *-ča* (= Gr. τε, Lat. *-que*, etc.).

Since the accent is not noted in writing in any Old or Middle Iranian language, its position and nature can only be deduced – as in Germanic – from its observable effects. In Avestan the most important phonological change connected with the accent is the devoicing of *r* (and *ər* < *\*r*) before *k*, *p*, *t*, which is restricted to forms in which the accent falls on the syllable containing *r*. The working of this rule, which results in written *hrk*,

*hrp* and (\*hrt >) *š*, indicates the existence of a free accent, which is often though not always on the same syllable as in the equivalent Vedic form, e.g. *vāhrka-* ‘wolf’, *amāša-* ‘immortal’ = Ved. *vā́rka-*, *amā́ta-*, but *mahrka-* ‘destruction’ = \*márka- (as against Ved. *marká-*). The formation of a compound or the addition of a suffix or enclitic (cf. the preceding paragraph) can result in a shift of accent, as in *amā́rta-tāt-* ‘immortality’ (cf. Ved. *sarvá-tāt(i)-* beside *sárva-*). See Mayrhofer in Schmitt 1989: 12–13, Beekes 1988: 55–69.

Whether the Avestan accent was still a musical (pitch) accent like that of Greek and Vedic or a dynamic (stress) accent is controversial, but there is no doubt that most Middle and Modern Iranian languages have developed a strong stress accent, which often causes syncope in unstressed syllables. In many Iranian languages the position of the stress has come to be wholly determined by the quantitative shape of the word, but a free stress, possibly reflecting the PIE accent, is still found in some modern Eastern Iranian languages; cf. Morgenstierne 1973a on the difference in stress in such pairs as Pashto *wúča* (f.) ‘dry’ (= Ved. *súšká-*) and *ričá* ‘nit’ (= Ved. *likšá-*).

The assumed close relationship between accent and ablaut (cf. above, p. 68) has become effaced in Iranian, as in other branches of PIE, to the extent that the accent can fall on any syllable, regardless of its vocalism. As a result of the merger of \**ě* and \**ō* in Ilr. *ǎ*, the PIE qualitative ablaut has disappeared, although the palatalization of the IE (labio-)velars before \**ě* occasionally allows its former presence to be discerned, as in the inflection of GAv. *aogah-* (n.) ‘strength’, acc. sg. *aogō*, instr. sg. *aojanhā* < \**h<sub>2</sub>ewg-os*, \**h<sub>2</sub>ewg-es-eh<sub>1</sub>*, cf. Gr. μένος, μένεος (Hoffmann 1958: 14–15), or the interrogative pronoun Av. *ka-*, *ča-* < \**k<sup>w</sup>o-*, \**k<sup>w</sup>e-* (cf. below, p. 277). Some such contrasts between forms with and without palatalization survive into Middle Iranian, as in Parthian *paryōž* beside *paryōγ* ‘victory’ or Khot. *tcamāna*, instr. sg. of *kye* ‘who’. On the other hand, the quantitative ablaut (the PIE alternation 0 ~ *e/o* ~ *ē/ō*) is well preserved and productive in Indo-Iranian, where it appears as 0 ~ *a* ~ *ā*, or, in combination with a following semi-vowel or consonant, *i/y* ~ *ai/ay* ~ *āi/āy*, *ɣ/r* ~ *ar* ~ *ār*, *a/n* (< \**n̥/n*) ~ *an* ~ *ān*, etc. It is to be noted that the Indo-Iranian long grade (*ā*, etc.) does not always derive from a PIE long grade but can also represent the o-grade by Brugmann’s Law (cf. above, p. 267).

These alternations, which can occur in any part of a word (root, suffix or ending), are of great importance for the historical morphology of Iranian (cf. also p. 282 below on the function of “*vṛddhi*” in word-formation). Ablaut occurs both within a single paradigm, a particular grade of the root and/or suffix being associated with each individual ending, and between contrasting paradigms.

Ablaut of the root is most often attested in formations without a suffix, particularly in root-presents such as *ah-/h-* ‘to be’ and reduplicated presents such as *dadā-/dad-* ‘to give’ (< \**de-deh<sub>3</sub>-/\*de-dh<sub>3</sub>-*). In formations containing a suffix (or infix) it is usually this element which shows alternation, e.g. nouns in *-tār-/tar-/tṛ-* (*-θr-*), athematic optatives in *-yā-/ī-* or present stems with infix *-na-/n-*. The preservation of an alternation in both root and suffix, as in Av. nom. sg. *pañtā*, gen. sg. *paθō* ‘path’ (< PIE \**pent-oh<sub>2</sub>-s/\*pñt-h<sub>2</sub>-es*), is exceptional.

Each individual form in such an alternating paradigm is characterized by a particular ablaut grade of the stem as well as by a specific ending. In the root-present, for instance, the 1/2/3 sg. pres. ind. active generally require the full grade of the stem (as in GAv. *mrao-mī*, etc., from *mrauu* ‘to say’), while the equivalent middle forms require the zero grade (°*mru-yē*, etc.). Similarly, a *u*-stem such as OPers. *Kuru-* ‘Cyrus’ has the zero grade of the stem in the nom. sg. (*Kur-u-š*) but the full grade in the gen. sg. (*Kur-au-š*).

Occasionally the occurrence of an abnormal ablaut grade (e.g. the long grade of the root in GAv. *stāumī*, 1 sg. pres. ind. act. of *stauu* ‘to praise’, or the long grade of the suffix in OPers. *dahəyāuš*, nom. sg. of the *u*-stem *dahəyu-* ‘country’) indicates that a category such as “root-present” or “*u*-stem” is not unitary but is made up of stems which originally belonged to various classes characterized by different configurations of accent and ablaut.

The endings do not normally display ablaut variation within a single paradigm, but only between contrasting paradigms (but cf. Table 4.22 on the inflection of Av. *xratu-*). Thus, the gen. sg. ending is attested as \*-as (PIE \*-es/-os) in Av. *rāiiiō*, *uxšnō* and OPers. *piça* (< \*piθras, cf. Gr. πατήρ, Lat. *patris*) from the stems *rāiii-* ‘wealth’, *uxšan-* ‘bull’ and *pitar-* ‘father’, but as \*-s in Av. *garōiš*, GAv. *čašmāng* (with *-ng* < \*-nh < \*-ns) and *nəṛəš* from the stems *gāri-* ‘mountain’, *čašman-* ‘eye’ and *nar-* ‘man’. Not all of the individual forms attested are ancient: *nəṛəš*, for instance, with its remarkable combination of the zero grade in both stem and ending, is probably an innovation for expected \*narō (cf. OInd. *nāras*, Gr. ἀνδρός). Nevertheless, since the innovation must have been based on an already existing form – in this case perhaps \*brā-tṛ-š (= OInd. *bhrātur*, ONor. *bróðor*), gen. sg. of *brātar-* ‘brother’ (cf. Hoffmann 1976: 598) – such a form can justifiably be used as evidence that Iranian inherited *r*-stems with “acrodynamic” accent and the associated type of ablaut (cf. above, p. 69).

## MORPHOLOGY

### Nouns

In Avestan, as in Old Indian, the system of three genders, three numbers and eight cases is well established (though it is only in the singular of a few declensions that all eight cases are formally distinct). During the later history of Iranian this system was gradually simplified. Old Persian has already reduced the cases to six by conflating the dative with the genitive and the instrumental with the ablative; Khotanese has gone further, retaining only remnants of the neuter gender and the dual number, while Sogdian has replaced most of the old plural inflections by forms derived from a collective noun in \*-tā-. Many Modern Iranian languages have dispensed both with the case system and with grammatical gender, so that in New Persian, for instance, the only morpheme surviving from the Old Iranian system of nominal inflection is the plural in *-ān* (< OPers. gen. pl. *-ānām*).

In Old Iranian the various declensions are principally distinguished by the final sound of the stem: stems in *a*, *ā*, *i*, *r*, etc. They are further divided into sub-classes by gender (e.g. stems in *a* into masculines and neuters) and, to a limited extent, by the different accent and ablaut patterns referred to on p. 273–274 above. The number of distinct declensions is very much reduced in Middle Iranian, where there is a marked tendency to transfer all masculine and neuter nouns to the *a*-declension and feminines to the *ā*-declension.

The most common declension in all Iranian languages is that of the masculine *a*-stems (PIE \*o-stems), whose inflection in Avestan (exemplified by *yasna-* ‘sacrifice, worship’ = Ved. *yajñá-*), Old Persian, Khotanese and Sogdian is shown in Table 4.21, together with the corresponding Vedic forms. (Only a selection of the numerous variant forms attested, especially in Avestan and Khotanese, is included in the table.)

TABLE 4.21 DECLENSION OF MASCULINE A-STEMS (PIE STEMS IN \*O)

	Vedic	Avestan	Old Persian	Khotanese	Sogdian
Sg. nom.	<i>yajñ-ás</i>	<i>yasn-ō</i>	<i>-a</i>	<i>-ā</i>	<i>-i</i>
acc.	<i>yajñ-ám</i>	<i>yasn-əm</i>	<i>-am</i>	<i>-u</i>	<i>-u</i>
instr.	<i>yajñ-ā</i>	<i>yasn-a</i>	<i>-ā</i>	<i>-na</i>	(= abl.)
dat.	<i>yajñ-āya</i>	<i>yasn-āi, GAv. also -āi.ā</i>	(= gen.)	(= gen.)	(= gen.)
abl.	<i>yajñ-āt</i>	<i>yasn-āt, LAv. also -āda</i>	<i>-ā</i>	(= instr.)	<i>-a</i>
gen.	<i>yajñ-āsya</i>	<i>yasn-ahe, GAv. -ahiā</i>	<i>-ahəyā</i>	<i>-i</i>	<i>-e</i>
loc.	<i>yajñ-é</i>	<i>yesn-e, yasn-aīia</i>	<i>-ai, -ayā</i>	<i>-a</i>	<i>-ya</i>
voc.	<i>yājñ-a</i>	<i>yasn-a</i>	<i>-ā</i>	<i>-a</i>	<i>-a</i>
Du. nom. acc.	<i>yajñ-ā, -āu</i>	<i>yasn-a</i>	<i>-ā</i>		<i>-a</i>
instr. dat. abl.	<i>yajñ-ābhyām</i>	<i>yasn-aēbiia</i>	<i>-aibiyā</i>		
gen.	<i>yajñ-āyoṣ</i>	<i>yasn-aīiā</i>			
loc.	(= gen.)	<i>yasn-aīiō</i>			
voc.	<i>yājñ-ā, -au</i>	<i>yasn-a</i>			
Pl. nom.	<i>yajñ-ās, -āsas</i>	<i>yasn-a</i>	<i>-ā</i>	<i>-a</i>	<i>-a</i>
acc.	<i>yajñ-ān</i>	<i>yasn-a, GAv. -əng</i>	(= nom.)	(= nom.)	
instr.	<i>yajñ-āis, -ébhiṣ</i>	<i>yasn-āiš</i>	<i>-aibiš</i>	<i>-yau</i>	
dat.	<i>yajñ-ébhyas</i>	<i>yasn-aēbiiō</i>	(= gen.)	(= gen.)	
abl.	(= dat.)	(= dat.)	(= dat.)	(= instr.)	
gen.	<i>yajñ-ānām</i>	<i>yasn-anəm</i>	<i>-ānām</i>	<i>-ānu</i>	<i>-ān</i>
loc.	<i>yajñ-ēsu</i>	<i>yasn-aēšu, -aēšuuu</i>	<i>-aišuvā</i>	<i>-uvo'</i>	
voc.	<i>yājñ-ās, -āsas</i>	<i>yasn-a</i>		(= instr.)	

This type of stem seems always to have had a fixed accent (with the exception of the voc. forms, which in Vedic are either unaccented – cf. Sogd. encl. voc. sg. *βay* ‘sir!’ beside stressed *βayá* – or accented on the first syllable regardless of the position of the accent in the rest of the paradigm, a rule for which there is some evidence also in Avestan; see Hoffmann 1975: 266 and cf. Gr. ἄδελφε ~ ἀδελφός, etc.). As for the individual endings, the majority of the Iranian forms are directly comparable with their Old Indian equivalents (see Table 4.10 above). The Av./OPers. instr. sg. in *-ā* corresponds to the rarer Ved. instr. in *-ā* rather than to that in *-ena* (which is of pronominal origin, as is Khot. *-na* < Old Iranian *\*-anā*). The usual Av. dat. sg. *-āi*, which may be compared directly with Gr. *-φ*, is more archaic than GAv. *-āi.ā*, OInd. *-āya*; the final *-ā* of the latter form seems to be a fossilized postposition, which may be found also in some Iran. abl. sg., loc. sg. and loc. pl. forms. In the nom./acc./voc. du. the Iranian forms agree with Ved. *-ā* (= Gr. *-ω*, cf. also Lat. *ambo*) rather than Ved. *-au*; the two forms are thought to be old sandhi-variants. In the remaining cases of the dual the Iranian and Indian forms are not precisely comparable, the most important difference being the preservation of a distinction between gen. and loc. du. in Avestan. In the nom./voc. pl. the regular equivalents of OInd. *-ās* and *-āsas* are the rare endings Av. *-ā* and *-āṅhō*, OPers. *-āha*, which seem to be particularly favoured for words pertaining to the sacral sphere (Av. *aməšā* ‘the immortal ones’, *yazatāṅhō* ‘(beings) worthy of worship’, OPers. *bagāha* ‘gods’). The usual form in both Avestan and Khotanese is *-a*, which has been explained as a PIE collective in *\*-ā* (< *\*-eh₂*), cf. Lat. *loca* ~ *locus* (Hoffmann 1958: 13); OPers. *-ā* and Sogd. *-a* are ambiguous and may equally well derive from *\*-ā* or *\*-ās* (or both).

Although Iranian inherited many varieties of stems showing ablaut variation (originally associated with a mobile accent), these seldom survive as independent types. As a result of a tendency to harmonize the inflection of all stems ending in the same sound (e.g. all stems in *u*), forms deriving from different ablaut types may be combined in the inflection of a single word, often making it difficult to discern its original ablaut pattern.



This point may be illustrated by the *u*-stem Av. *xratu-* (m.) ‘mental power, intention, etc.’ (= OInd. *krātu-* ‘power’), of which all the attested forms are shown in Table 4.22. (The only forms which occur in Old Persian are the two acc. sg. forms *xratum* and *xraθum*, the latter showing generalization of *θ* from a form such as instr. sg. \**xraθuvā* = Av. *xraθβā*.)

TABLE 4.22 DECLENSION OF AV. *xratu-*

	sg. (GAv.)	(LAv.)	pl. (GAv.)
nom.	<i>xratuš</i>	<i>xratuš</i>	<i>xratauuō</i>
acc.	<i>xratūm</i>	<i>xratūm, xraθβam</i>	<i>xratūš</i>
instr.	<i>xratū, xraθβā</i>	<i>xraθβa</i>	* <i>xratubīš</i>
dat.		<i>xraθβe</i>	* <i>xratubiō</i>
abl.	(= gen.)	<i>xrataoŋ</i>	(= dat.)
gen.	<i>xratəuš</i>	<i>xratəuš, xraθβō</i>	* <i>xratunəm</i>
loc.	<i>xratā</i>		* <i>xratušū</i>
voc.		<i>[hu]xratuuō</i>	(= nom.)

In this paradigm the suffix appears in the zero grade as \**u* (Av. *ū*) or \**w* (> Av. *β* after *θ*), in the full grade as \**au* (Av. *əu/ao*, in final position *ō/uuō*) or \**aw* (Av. *auu*) and in the long grade as \**āu* (> Av. *ā* in final position). Note also the occurrence of two ablaut variants of the ending itself in the instr. sg. (*-ū* < \**-u-h<sub>1</sub>*; *-βā* < \**-w-eh<sub>1</sub>* or \**-w-oh<sub>1</sub>*) and gen. sg. (*-əuš* < \**-ow-s*; *-βō* < \**-w-es* or \**-w-os*). The etymologies of the remaining endings are as follows. Singular: nom. *-š* < \**-s*; acc. *-m* < \**-m* (LAv. variant *-əm* borrowed from the *a*-stems); dat. *-ē* < \**-ey*. Originally the abl. sg. was formally distinct from the gen. only in the *a*-declension; LAv. *xrataoŋ* exemplifies a later tendency to create special abl. forms by borrowing the final *-ŋ* of the *a*-stems. The loc. and voc. sg. are both endingless but differ in the grade of the suffix. Plural: nom./voc. *-ō* < \**-es*; acc. *-(ū)š* < \**-(u-)ns*; instr. *-bīš* < \**-b<sup>h</sup>is*; dat./abl. *-biiō* < \**-b<sup>h</sup>iyos*; loc. *-šū* < \**-su*. The gen. pl. (like that of the *a*-stems and most other declensions) was remodelled in Indo-Iranian after that of the *n*-stems, but the older ending *-əm* (< \**-ōm*, Gr. *-ov*) is occasionally attested, as in LAv. *yāθβəm* (beside *yātuəm*), gen. pl. of *yātu-* ‘sorcerer’.

### Adjectives, pronouns and numerals

In general, adjectives are inflected exactly like nouns, though a few common adjectives, such as Av. *vīspa-* ‘all’ and its cognates, display some of the peculiarities of pronominal declension (see below), e.g. Sogd. abl. (originally instr.) sg. m. *wispna*, LAv. nom. pl. m. *vīspne* (= Khot. *biššā*, contrast with GAv. *vīspāhō*), gen. pl. m. *vīspaēšəm* (beside *vīspanəm*). The feminine forms of adjectives are usually derived from a separate stem in *-ā* or *-ī* (even where the m./n. stem belongs to a class, such as the *u*-declension, which includes feminine nouns). Examples from Avestan: *sūra-*, f. *sūrā-* ‘strong’; *po<sup>o</sup>ru-*, f. *pa<sup>o</sup>rī-* ‘much’; *bə<sup>o</sup>zant-*, f. *bə<sup>o</sup>zā<sup>o</sup>tī-* ‘high’.

As in Old Indian, comparatives and superlatives can be formed in two ways: with the suffixes *-tara-* and *-tama-* added to the stem of the positive (e.g. Av. *aš.aojah-*, *aš.aojastara-*, *aš.aojas-tama-* ‘possessing much, more, most power’) or with the suffixes *-yah-* and *-išta-* added directly to the underlying root in the full grade (e.g. Av. *uγ-ra-*, *aoj-īah-*, *aoj-išta-* ‘strong, -er, -est’). Also formed directly from the root is the compound form in *-i-*, as in Av. *tiži.asūra-* ‘sharp-tusked’ (< \**tij-i-* beside *tij-ra-* ‘sharp’), *bə<sup>o</sup>zi.čaxra-* ‘high-wheeled’ (beside *bə<sup>o</sup>zant-*), cf. OInd. *ṛj-i-*, Gr. *ἄργι-ι-* as the compound form of *ṛj-rá-*, *ἄργός* (< \**ἄργι-pó-ς*) ‘swift; bright’, etc. (cf. above, p. 79).

The principal Avestan demonstrative pronouns are *hō* (nom. sg. m.), *hā* (nom. sg. f.), *taṭ* (nom. sg. n.) ‘this; he, she, it’, and its compound *aēšō*, *aēša*, *aētaṭ*; *aēm*, *īm*, *imaṭ* ‘this’; and *hāu*, *hāu*, *auuaṭ* ‘that’. In their inflection these show the same kinds of peculiarities as the equivalent Old Indian forms (cf. above, p. 85 ff.), including the employment of suppletive stems, often opposing the nom. sg. m. and f. (e.g. *hō*, *hā*) to the rest of the declension (stem *ta-*), and the prefixation or suffixation of deictic particles (e.g. *aē-* in *aē-ša-*, etc., \*-am in *aēm*, *īm* = OInd. *ay-ám*, *iy-ám*). The use of certain endings different from those of nouns (e.g. nom./acc. sg. n. in *-t*, instr. sg. m./n. in *-na*, nom. pl. m. in *-e*) and the infixation of additional elements between the stem and ending (e.g. *-hm-* and *-hy-* respectively in several cases of the m. and f. sg., *-h/-š-* in the gen. pl.) may be exemplified by the following forms of the demonstrative Av. *aēm* ‘this’ (stems *ay-/i-*, *a-* and *ima-*): nom. sg. n. *ima-t*, instr. sg. m. *a-na*, dat. sg. m. *a-hm-āi*, dat. sg. f. *a-ḥh-āi* (< \*a-hy-āi), nom. pl. m. *im-e*, gen. pl. m. *aē-š-qm*, gen. pl. f. *ā-ḥh-qm*. The Old Persian forms follow the same principles.

Similar irregularities occur in the inflection of the relative pronoun, Av. *yō* (GAv. *yō*), *yā*, *yaṭ*, OPers. *haya*, *hayā*, *taya* (where the relative has been compounded with the demonstrative \*hā-, \*ta-), and of the interrogative pronouns. In Old Iranian, unlike Old Indian, all of the four interrogative stems, *ka-*, *kā-*, *ča-* and *či-*, still function as pronouns and tend to combine into a suppletive system like that of the demonstratives: Av. *kō*, *kā*, *čūt* (nom. sg. m., f., n.), cf. OPers. *kaš-či* ‘someone’, *čiš-či* ‘something’.

The inflection of the personal pronouns differs even more markedly from that of nouns, as may be illustrated by the following selection of first person forms:

nom. sg.	Av. <i>azəm</i> , OPers. <i>adam</i>
acc. sg.	LAv. <i>maqm</i> , OPers. <i>mām</i>
dat. sg.	GAv. <i>ma<sup>i</sup>biiā</i> , <i>ma<sup>i</sup>biiō</i> , LAv. <i>māuur<sup>o</sup>iia</i>
gen. sg.	LAv. <i>mana</i> , OPers. <i>manā</i>
nom. pl.	Av. <i>vaēm</i> , OPers. <i>vayam</i>
dat. pl.	GAv. <i>ahma<sup>i</sup>biiā</i>
gen. pl.	LAv. <i>ahmākəm</i> , OPers. <i>amāxam</i>

(cf. OInd. *ahám*, *mām*, *máhya(m)*, *máma*; *vayám*, *asmábhya(m)*, *asmákam*).

It will be noted that, as in Old Indian, these forms show no distinction of gender and that the singular and plural forms are derived from apparently unrelated stems. A further peculiarity of the personal pronouns is the existence of alternative unaccented (enclitic) forms in certain cases, e.g. 1 sg. Av. *mā*, OPers. *-mā* (acc.), GAv. *mōi*, LAv. *mē*, OPers. *-mai* (gen./dat.). In the plural, Old Avestan preserves a distinction between the encl. acc. *nā* ‘us’, *vā* ‘you’ (cf. Lat. *nōs*, *vōs*) and the encl. gen./dat. *nā*, *vā*, while LAv. *nō* and *vō*, like OInd. *nas* and *vas*, are used for all three cases. Finally, we may note the GAv. nom. sg. f. forms *θβōi* and *x<sup>a</sup>aē[čā]* (from the possessive adjectives *θβā-* ‘thy’, *x<sup>a</sup>ā-* ‘(one’s) own’), whose ending may be compared with that of Lat. *quae*, etc. (Hoffmann 1958: 16).

The numerals show many peculiarities, both archaisms and innovations, in their inflection and word-formation. These are discussed in detail by Emmerick (1992), whose study shows that the rich data from Middle and Modern Iranian languages can be useful in clarifying the Old Iranian forms.

## Verbs

In Old Iranian, and especially in Avestan, the inflection of the verb is extremely rich as a result of the numerous intersecting categories into which its forms are classified: person (first, second or third), number (singular, dual or plural), mood (indicative, injunctive, subjunctive, optative or imperative), tense (present, aorist, perfect, etc.) and voice

(active, middle or passive). In general, the category of tense is indicated by the stem of the verb; that of mood by the presence or absence of a modal suffix following the tense-stem, the presence or absence of the augment *a-* before the tense-stem and the choice of ending; those of person, number and voice by the verbal endings alone (except in the case of the passive present stem in *-ya-*). In addition to the finite forms of the verb, participles and infinitives are also attested. The following survey (based on the comprehensive description of Kellens 1984) is primarily concerned with Avestan; Old Persian provides examples of most of the corresponding types and categories but no complete paradigms.

Present stems can be formed in many ways, of which only the principal types can be mentioned here. The most important division is that between “thematic” and “athematic” presents. The thematic presents are formed by adding to the root (in a particular, invariably ablaut grade) a suffix consisting of or ending in *-a-* < PIE *\*-e/o-*:

Av. *θβər³s-a-* ‘to fashion’ (zero grade of root + suffix *-a-*)  
*bauu-a-* ‘to become’ (full grade + *-a-*)  
*bīδ-īia-* ‘to notice’ (zero grade + *-ya-*)  
*zb-aiia-* ‘to invoke’ (zero grade + *-aya-*)  
*baṇd-aiia-* ‘to bind’ (full grade + *-aya-*)  
*xšnāuu-aiia-* ‘to satisfy’ (long grade + *-aya-*)  
*ja-sa-* ‘to come, go’ (zero grade + *-sa-* < PIE ‘inchoative’ *\*-sko-*)

The various types of athematic present have in common certain endings different from those of thematic stems (see below on the indicative and imperative) and the occurrence of ablaut alternation in the stem:

*jan-/yn-* ‘to strike’ (root-present)  
*da-dā-/da-d-* ‘to give’ (reduplicated present)  
*vi-na-d-/vi-ṇ-d-* ‘to find’ (infix nasal)  
*d³b²-nao-/d³b²-nu-* ‘to deceive’ (zero grade + *-nao-/nu-*)  
*stər²-nā-/stər²-n-* ‘to spread’ (zero grade + *-nā-/n-*)

In origin, these last two classes are special cases of the preceding type, the nasal infix having been inserted into a root with final *\*-w-* or *\*-H-*, cf. *ā.d³bao-man-* ‘deception’ (which demonstrates the existence of a root *dbav* beside *dab*), *star²ta-* ‘spread’ (< *\*stṛh₃-to-*, Gr. *σπρωτός*), etc.

Certain types of present stem, notably the passives in *-ya-*, causatives in *-aya-* and future stems in *\*-sya-* (> *-hya-*, *-šya-*), express a special or modified sense of the verb:

*jan-īia-* ‘to be struck’ (beside *jan-/yn-* ‘to strike’)  
*jām-aiia-* ‘to cause to go’ (beside *ja-sa-* ‘to come, go’)  
 fut. ptcp. act. *bū-šīia-ṇt-* ‘about to be’ (beside pres. ptcp. act. *bauu-a-ṇt-* ‘being’)

The future stem is most often attested by its participles, the sense of a fut. indicative being more commonly expressed by the pres. subjunctive. The passive stem in *-ya-* (which in Iranian, unlike Old Indian, takes mid. or act. endings indifferently) is also comparatively rare, in part as a result of the fact that a passive sense can alternatively be expressed by the use of the normal (non-passive) pres. stem with mid. instead of act. endings, e.g. *vaēna'te* (mid.) ‘is seen, seems’ as opposed to *vaēna'ti* (act.) ‘sees’.

The principal types of aorist stem are the sigmatic aorist, e.g. *xšnāu-š-/xšnao-š-* ‘to satisfy’, in which the suffix *\*-s-* (> *-s-*, *-h-*, *-š-* according to phonological context) is combined with alternation between the long grade and full grade of the root, and the root-aorist, e.g. *jam-/ym-* ‘to come, go’, which displays alternation between the full grade and zero grade as in the most common type of root-present (though the distribution of the two alternants is slightly different in the aorist). The perfect stem is usually formed by reduplication, e.g. *va-uuac-/va-oc-* ‘to say’. As in other IE languages, the verb ‘to know’ irregularly forms an unreduplicated perf. stem *vaēd-/vid-*, cf. OInd. *vēda*, *vidmā*, Gr. (ϕ)οἶδα, (ϕ)ῖδμεν (cf. above, p. 98), etc. The role of the aor. and perf. stems is very much restricted in Later Avestan – even more so in Old Persian – a development marking the first stage in the creation of the Middle Iranian verbal system (based on the pres. stem and a new “past stem” derived from the past participle in *-ta-*).

It is convenient to begin a survey of the formation of the moods with the injunctive, which is formed by the addition of the so-called secondary endings – in fact the verbal endings in their most basic forms, see above, p. 95 – directly to the pres. or aor. stem. The “secondary” endings (omitting those of the dual, since they are poorly attested and often etymologically obscure) are given in Table 4.23. With the exception of 2 sg. mid. *\*-sa* (cf. Gr. ἔθου, Hom. ἔθεο < *\*h<sub>1</sub>e-d<sup>h</sup>h<sub>1</sub>-so*, etc., as against OInd. *-thās*) and 3 pl. act. *-at* (< *\*-nt*), an ablaut variant lost in Old Indian, these endings correspond precisely to the equivalent Old Indian forms.

TABLE 4.23 SECONDARY ENDINGS IN AVESTAN

		active	middle
sg.	1	<i>-m</i> or <i>-əm</i>	<i>-i</i>
	2	<i>*-s</i> (> <i>-h-</i> , <i>-š-</i> , etc.)	<i>*-sa</i> (> <i>-ha-</i> , <i>-ša-</i> , etc.)
	3	<i>-t</i>	<i>-ta</i>
pl.	1	<i>-ma</i>	<i>-ma<sup>h</sup>di</i>
	2	<i>-ta</i>	<i>-dām</i> or <i>-dβām</i>
	3	<i>-ən</i> (< <i>*-e/ont</i> ), <i>-n</i> (< <i>*-nt</i> ) or <i>-at</i> (< <i>*-nt</i> )	<i>-nta</i> or <i>-ata</i> (< <i>*-nta</i> )

The imperfect is formed, as in Old Indian, by prefixing the augment *a-* (= OInd. *a-*, Gr. ἔ-, Arm. *e-*) to the pres. injunctive. The imperfect is well attested in Old Persian and in some later Iranian languages such as Sogdian, but comparatively rare in Avestan, where the pres. injunctive has largely taken over its function as a past tense. The even rarer aor. indicative, of which a few forms are attested in Old Avestan and Old Persian, is similarly formed by the prefixation of the augment to the aor. injunctive. The pres. and perf. indicatives, however, are characterized in a different way, by the use of endings distinct from those of the injunctive.

The so-called primary endings of the pres. indicative (again omitting the dual forms), as attached to thematic pres. stems, are given in Table 4.24. All of these endings have exact cognates in Old Indian. The inflection of thematic stems differs only in the 1 sg. active, where Old Avestan attests the ending *-ā* (= Gr. *-ω*, Lat. *-ō*, etc.) as against OPers./LAv./OInd. *-ā-mi*. The thematic vowel, in general *a*, appears as *ā* (< *\*o* by Brugmann’s Law, cf. above, p. 267) in 1 pl. act. *-ā-mahi* and mid. *-ā-ma<sup>h</sup>de*; on the other hand, the thematic 1 sg. middle has merely *-e* < *\*-ai*, where *\*-āi* < *\*-a-ai* might have been expected. Since the thematic pres. indicative is well attested in most Iranian languages, it is possible to give some complete paradigms, at least of the sg. and pl. forms, based on the pres. indicative of *bar*, pres. stem *bar-a-* (act.) ‘to carry’, (mid.) ‘to ride’; see Tables 4.25 and 4.26.

TABLE 4.24 PRIMARY ENDINGS IN AVESTAN

		active	middle
sg.	1	-mi	-e
	2	*-si (> -hi, -ši, etc.)	*-sai (> -he, -še, etc.)
	3	-ti	-te
pl.	1	-mahi	-ma'de
	2	*-θa	-duiiē
	3	-ənti, -nti or -a'ti (< *-nti)	-nte or -a'te (< *-ntoy)

TABLE 4.25 CONJUGATION OF THE THEMATIC PRESENT INDICATIVE ACTIVE

		Avestan	Old Persian	Khotanese	Sogdian
sg.	1	<i>bar-ā-mi</i>	<i>bar-ā-mi</i>	<i>barāmā</i>	<i>βarām</i>
	2	<i>bar-a-hi</i>		<i>bīri</i>	<i>βare</i>
	3	<i>bar-a-ti</i>	<i>bar-a-ti</i>	<i>bīdā</i>	<i>βarti/βart</i>
pl.	1	<i>bar-ā-mahi</i>	<i>bar-ā-mahi</i>	<i>barāmā</i>	<i>βarēm</i>
	2	<i>bar-a-θa</i>		<i>baḍa</i>	<i>βarθa/βarta</i>
	3	<i>bar-ə-nti</i>	<i>bar-a-nti</i>	<i>barīndā</i>	<i>βarand</i>

TABLE 4.26 CONJUGATION OF THE THEMATIC PRESENT INDICATIVE MIDDLE

		Avestan	Old Persian	Khotanese
sg.	1	<i>ba'r-e</i>	<i>bar-ai</i>	<i>bare</i>
	2	<i>bar-a-he</i>		<i>bara</i>
	3	<i>bar-a-te</i>	<i>bar-a-tai</i>	<i>baḍe</i>
pl.	1	<i>bar-ā-ma'de</i>		<i>barāmāne</i>
	2	(GAv.) <i>bar-a-duiiē</i>		<i>barīru</i>
	3	<i>bar-ə-nte</i>		<i>barāre</i>

The 2 pl. act. ending \*-ta in Khot. *baḍa*, Sogd. *βarta* (beside *βarθa*) is a secondary ending borrowed from the injunctive, etc., as is the 2 sg. mid. \*-ha in Khot. *bara*; the Khot. 2 pl. mid. ending *-īru* is borrowed from the optative. The ending of Khot. 1 pl. act. *barāmā* seems to correspond more closely with Classical Skr. *-mas* than with its Vedic variant *-masi* (= OPers./Av. *-mahi*). A more significant division amongst the various Iranian languages is found in the 3 pl. middle, where Khotanese and some other languages attest an ending \*-ārai (= Av. *-ā're*) or \*-rai (= OInd./Av. *-re*). In Avestan, as in Old Indian, this ending is restricted to a small group of root-presents, some of which also have a 3 sg. middle in *-e* rather than *-te*. These special endings, which are also found in the perf. ind. middle, seem originally to have characterized a particular sub-class of root-presents (with a fixed accent on the root and ablaut alternation between long grade and full grade instead of between full grade and zero grade; see Narten 1968).

The endings of the perf. ind. active (sg. and pl.) are as follows: singular: 1 *-a*, 2 *-θa*, 3 *-a*; plural: 1 *-ma*, 3 *-ar<sup>ə</sup>* or *-ar<sup>ə</sup>š*. It is not clear which of the two 3 pl. endings is to be equated with OInd. *-ur* (< \*-r̥ = *-ar<sup>ə</sup>* or < \*-r̥s = *-ar<sup>ə</sup>š*). The endings of the perf. ind. middle, in so far as they are attested, are the same as those of the pres. ind. middle, with 3 sg. *-e* and probably 3 pl. \*-re (cf. Khot. *byaure* 'they exist' < \*abi-āf-rai, originally 3 pl. perf. middle of *abi-āp* 'to find, obtain'); see above.

The subjunctive is characterized by a suffix *-a-*, which is inserted between the stem (whether present, aorist or perfect) and the endings. In the case of thematic stems, the

subj. suffix combines with the final vowel of the stem to a long  $\bar{a}$ . The endings are a mixture of primary and secondary, the choice being fixed in some cases and free in others, except in the first person singular, where GAv. act.  $-\bar{a}$  and mid.  $-\bar{a}i$  are later replaced by the special endings  $-\bar{a}ni$  (= OInd.  $-\bar{a}ni$  beside  $-\bar{a}$ ; cf. above, Table 4.13) and  $-\bar{a}ne$  respectively.

The optative is similarly characterized by the insertion of a suffix between the pres., aor. or perf. stem and the endings, which in this case are always the secondary endings, apart from the special endings 3 pl. act.  $-\bar{a}r^o$  or  $-\bar{a}r^o\check{s}$  (beside secondary  $-n$ ) and 1 sg. mid.  $-a$ . In the case of most athematic stems, the optative suffix shows ablaut alternation between  $-y\bar{a}-$  and  $-\bar{i}-$  (from  $*-yeh_1-/*-ih_1-$ ). In all other cases the suffix is a non-alternating  $-\bar{i}-$ , which combines with the final vowel of thematic stems to form the diphthong  $*ai$  (> Av.  $a\bar{e}$  or  $\bar{o}i$ ). A special feature of Iranian (attested in Avestan, Old Persian and Sogdian) is the employment of the augment with certain optative forms which express a repeated or habitual action in the past (cf. below, p. 282).

The endings of the imperative are added directly to the pres. or aor. stem. (No perf. imperative is attested.) Active: singular: 2 (thematic)  $-0$ , (athematic)  $-di$ , 3  $-tu$ ; plural: 2  $-ta$ , 3  $-\bar{e}ntu$  or  $-\ntu$ . Middle: singular: 2  $*-swa$  (>  $-suua$ ,  $-huua$ ,  $-\check{s}uua$ ), 3  $-tqm$  or  $-qm$ ; plural: 2  $-d\bar{u}m$  or  $-\delta\beta\bar{a}m$ , 3  $-\ntqm$ . These endings, all of which have exact cognates in Old Indian, are peculiar to the imperative (except for those of the second person plural, which are identical with the secondary endings). There is no first person imperative in Iranian.

Pres. and aor. stems form their active participles by means of the suffix  $-\bar{a}nt-/-at-$  (athematic) or  $-\nt-$  (thematic), while perf. stems employ the suffix  $-uuah-/-u\check{s}$ . All three types of stem form their middle participles in the same way, with the suffix  $-\bar{a}na-$  (athematic) or  $-mna-$  (thematic). The latter form may be directly equated with Gr.  $-\mu\epsilon\nu\omicron\varsigma$  (<  $*-mh_1no-$ ), while its OInd. equivalent,  $-m\bar{a}na-$ , shows the influence of the athematic suffix  $-\bar{a}na-$  (<  $*-mh_1no-$ ). Certain other verbal adjectives or participles are not formed from a tense-stem but directly from the root, the most important being the ‘past participle’ in  $-ta-$ , which has a passive sense in the case of transitive verbs and which comes to provide the basis for all the past tense formations in most Middle and Modern Iranian languages. Various types of infinitive are attested in Avestan, although none of them are common. As in the case of the participles, some are derived from a tense-stem, others directly from the root. The Old Persian infinitives, on the other hand, are all of a single type (not found in Avestan or Old Indian), the suffix  $-tanai$  being added to the full grade (PIE e-grade) of the root, e.g.  $\check{c}artanai$  ‘to do’ from the root  $kar$ .

## WORD-FORMATION

The principal means of creating new nominal stems in Iranian are suffixation and compounding, the individual suffixes and types of compound being largely identical with those found in Old Indian. There are also a few prefixes, such as  $a-$ ,  $an-$  ‘un-’ and  $hu-$  ‘good’, cf. above, p. 79–80. A peculiarity of Avestan, of which traces survive in some Middle Iranian languages, is the tendency to replace the bare stem by the nom. sg. form, both in compounds and before certain suffixes, e.g.  $b\bar{a}zu\check{s}.ao\check{j}ah-$  ‘strong-armed’,  $da\bar{e}uu\bar{o}$ .  $d\bar{a}ta-$  ‘devil-created’,  $da\bar{e}uu\bar{o}.t\bar{a}ma-$  ‘arch-devil’ (beside  $b\bar{a}zu.stauuah-$  ‘as thick as an arm’,  $da\bar{e}uua-iiasna-$  ‘devil-worshipper’, etc.). A further Avestan development is the employment of the compound form in  $-\bar{o}^o$ , originally the nom. sg. m. of the  $a$ -declension (cf.  $da\bar{e}uu\bar{o}.d\bar{a}ta-$ , etc.), without regard to the gender or declension of the stem, as in  $da\bar{e}n\bar{o}.s\bar{a}\check{c}$ - ‘well versed in the religion’ (from the feminine noun  $da\bar{e}n\bar{a}$ -) or  $kar^ap\bar{o}.t\bar{a}t-$  (a collective noun derived from  $kar^apan-$ , the designation of a class of priests). Similarly, in Sogdian, a feminine  $\bar{a}$ -stem such as  $x\bar{a}n\bar{a}$  <  $*x\bar{a}n\bar{a}k\bar{a}$ - ‘house’ appears before certain

suffixes as *xānē-* < \**xānāki-* (-i being the nom. sg. m. ending of the Sogd. *a*-declension; cf. Table 4.21), e.g. pl. *xānē-t* ‘houses’, in origin a collective noun with suffix \*-tā-.

The use of “vṛddhi” of the first syllable as a derivational device (cf. above, p. 273) is well established in Iranian, although it never became common as it did in Classical Sanskrit. As parallels to the Old Indian forms with *ā* and *ār* as vṛddhi of *a* and *ṛ/ra* respectively, one may cite such forms as Av. *hāuuani-* ‘(time) appropriate for pressing’ from \**hauuana-* ‘act of pressing’ (OInd. *savana-*), *vār<sup>o</sup>θraṇi-* ‘victorious’ from *vār<sup>o</sup>θraṇa-* ‘victory’, OPers. *Mārgava-* ‘inhabitant of Margu’. Some Iranian languages seem to have agreed with Old Indian also in using the long diphthongs *āi*, *āu* (= OInd. *ai*, *au*) as vṛddhi of *i* and *u*; cf. the OPers. month-name *Θāigrači-*, probably from \**θigra-ka-* ‘garlic’ (cf. Pers. *sīr* ‘id.’ < \**θigra-*), MPers. *wāspuhr* (< \**wāispuhr*) ‘principal’ from *wispuhr* ‘prince’, etc. In such cases, however, Avestan consistently follows an older derivational pattern in employing the short diphthongs \**ai*, \**au* (> *aē*, *ao*, etc.), as in *duuaēpa-* ‘island’ (contrast OInd. *dvīpā-* ‘island’ < \**dwi-h<sub>2</sub>p-o-* ‘(between) two waters’, without vṛddhi), *daožan<sup>h</sup>a-* ‘hell’ from *duž-ahu-* ‘id.’ (literally ‘evil existence’).

## SYNTAX

Traditionally, much less attention has been paid to the syntax of the Iranian languages than to their phonology and morphology. In this respect many of the chapters in Windfuhr 2009 offer a refreshing break with tradition, dealing systematically with topics such as word-order, noun phrase structure, use of the cases, tenses and moods, co-ordination and subordination. On the syntax of Old Avestan see also West 2011. Here it must suffice to mention some of the more important points in which the Iranian languages differ from Old Indian.

One of the most remarkable features of Old Iranian nominal syntax is the ability of the instr. pl. form to substitute for other cases of the plural, as in Av. *vīspāiš aoi karšuuṇ yāiš hapta* ‘to all the seven continents’ (instr. *vīspāiš*, *yāiš* for acc.), OPers. *XIV raučabiš θakatā āha* ‘14 days had passed’ (instr. *raučabiš* for nom.). Cf. also the use of the instr. pl. for the voc. pl. in Khotanese (see Table 4.21) and as a generalized oblique case of the plural in some of the Modern Iranian languages of the Pamir mountains (Wakhi -*əv* < \*-*aibiš*, etc.). Equally noteworthy is the use of the relative pronoun (Av. *yā-*, OPers. *hayā-/taya-*, see p. 277 above) in attributive constructions such as Av. *daēūm yim apaošəm* ‘the demon Apaosha’ (acc.) or OPers. *dahəyūnām tayaišām parūnām* ‘of many lands’, a usage which results from the reinterpretation of a nominal relative clause such as Av. *daēuuō yō apaošō*, originally ‘the demon who (is) Apaosha’, and the attraction of the relative pronoun (and predicate) into the case of the antecedent, giving *daēūm yim apaošəm* for \**daēūm yō apaošō*. (See Reichelt 1909: 370–371.)

Several characteristic features of the syntax of the verb in Old Iranian have already been referred to above, including the use of the injunctive rather than the imperfect as the normal narrative past tense, which is peculiar to Avestan, and the use of the optative (sometimes with augment) to express a repeated or habitual action in the past, e.g. Av. *tūm zəmargūzō ākər<sup>o</sup>nūuō vīspe daēuua, zaraθuštra, yōi para ahmāt vīrō.raoða apataian pətī āiia z<sup>o</sup>mā* ‘you, Zarathushtra, drove underground all the demons who previously used to go about on this earth in human form’; OPers. *yaθā-šām hačā-ma aθaḥəya, avaθā akunavayantā* ‘as was said to them by me, so they used to do’; Sogd. *čāf awya nāra awī ḍasta nīyāse, ə<sup>o</sup>ti-šī xa nāra čan ḍasta wāpate* ‘however many pomegranates she took in (her) hands, the pomegranates fell from her hands’.

The loss of the PIE perfect system, which is incipient in Late Avestan and almost complete in Old Persian, is made good by the creation of a new type of perfect based on the past participle (with an obligatory passive construction, the agent being originally in the dative, replaced in Old Persian by the genitive): Av. *yezi-ča hē anīia aya šīiaoθna frauuaršta* ‘and if he has committed other evil deeds’; OPers. *ima taya manā kartam* ‘this (is) what I have done’, lit. ‘what (has been) done by me’. In many later Iranian languages this construction comes to express a simple past tense, as in MPers. *man kard* ‘I did (it)’. Another verbal periphrasis which later becomes widespread, especially in Eastern Middle Iranian, is the so-called potential construction (Sims-Williams 2007). This is first attested in Old Persian, where the past participle of a transitive verb is used with the auxiliary *kar* ‘to make’ (in the active) or *bav* ‘to become’ (in the passive) to express either a potentiality or the consummation of an action: *nai āha martiya . . . haya avam Gaumātam tayam magum xšaçam dītam čaxriyā* ‘there was no-one . . . who could have deprived that Gaumāta the magus of the kingship’; *yaθā kantam abava, pasāva θikā avaniya* ‘when it had been dug, then it was filled with gravel’. In Middle Iranian the potential construction also occurs with intransitive verbs (auxiliary ‘to become’, Sogd. *βw-*, Khot. *hām-*), e.g. Sogd. *ne nipasta βōt* ‘he cannot lie down’; Khot. *ku vā drai māštā parrāte hāmāte, balysā rrundu kšamotte* ‘when three months had passed, the Buddha took leave of the king’.

Note: The author would like to thank Almut Hintze and Philip Huysse for valuable comments on earlier drafts of this chapter.

## FURTHER READING

In addition to bibliographic details of works cited in this chapter the following list includes a selection of basic literature on the Iranian languages, especially Avestan and Old Persian.

The most up-to-date survey of the whole field, with a primarily typological rather than historical orientation, is to be found in Windfuhr 2009, with a monograph-length chapter on Old Iranian (p. 43–195) by P. O. Skjærvø. The *Compendium Linguarum Iranicarum* (Schmitt 1989), which includes chapters on the prehistory of the Iranian languages (p. 4–24, by M. Mayrhofer), on Old Persian (p. 56–85, by R. Schmitt) and on Avestan (p. 32–55, by J. Kellens), is still very useful. For Avestan see also K. Hoffmann 1987. An earlier article by the same author (Hoffmann 1958) gives an incisive characterization of the special features of Old Iranian, as compared with Old Indian, a topic covered in greater detail in Goto 2013. These surveys supplement but do not replace Geiger & Kuhn 1895–1903, whose chapters on the Old Iranian languages (p. 1–248, by Christian Bartholomae), though in many respects dated, remain unsurpassed in comprehensiveness.

An etymological dictionary of the Iranian languages (Rastorgueva and Èdel'man 2000–2007, Èdel'man 2011–, in Russian) is in progress; Cheung 2007 lists Iranian verbal roots with their derivatives.

The standard edition of the Avesta is that of K. F. Geldner (1886–1896, in the original script). The dictionary of Bartholomae (1904) has likewise not been superseded. Martínez and de Vaan 2014 provide an introduction to Avestan from an IE perspective and include text specimens, a glossary and an up-to-date basic bibliography. On Avestan syntax see Reichelt 1909: 218–387 and Skjærvø 2009: 94–166. A modern treatment of Avestan phonology and morphology is provided by Hoffmann and Forssman 2004; see also Kellens 1984 and 1995 on the morphology and syntax of the Avestan verb.



All the works mentioned above cover both Old and Later Avestan. Modern editions of the Old Avestan texts, with translation and commentary, include Insler 1975 (Gāthās only), Narten 1986 and Hintze 2007 (both Yasna Haptanhā'ti only), Kellens and Pirart 1988–91, Humbach 1991 and Humbach and Faiss 2010. The phonology and morphology of Old Avestan are treated in Beekes 1988 (cf. also Kellens & Pirart 1988: 42–88 on “phonétique et graphie”), aspects of its syntax in Kellens and Pirart 1990 (which also contains a complete lexicon to the Old Avestan texts) and in West 2011.

The most comprehensive edition of the Old Persian inscriptions is that of R. G. Kent (1953; supplemented by Mayrhofer 1978). Kent's book also contains a historical grammar (more detailed but less reliable than Schmitt 1989: 56–85 or de Vaan & Lubotsky 2012) and a lexicon, the latter now superseded by Schmitt 2014. Several of the most important inscriptions have been re-edited by R. Schmitt (1991, 2000).

Finally, the existence of useful on-line resources for a number of Iranian languages should not be overlooked. These include introductory texts by P. Oktor Skjærvø on Avestan, Old Persian and Sogdian (all at <http://www.fas.harvard.edu/~iranian/>).

## REFERENCES

- Bartholomae, Christian 1904, *Altiranisches Wörterbuch*, Trübner, Strassburg
- Beekes, Robert S. P. 1988, *A Grammar of Gatha-Avestan*, Brill, Leiden
- Cheung, Johnny 2007, *Etymological Dictionary of the Iranian Verb* (Leiden Indo-European Etymological Dictionary Series, 2), Brill, Leiden
- de Vaan, Michiel 2003, *The Avestan Vowels* (Leiden Studies in Indo-European, 12), Rodopi, Amsterdam
- de Vaan, Michiel & Lubotsky, Alexander 2012, “Old Persian,” in: Holger Gzella (ed.), *Languages from the World of the Bible*, de Gruyter, Berlin: 194–208
- Degener, Almuth 2002, “The Nuristani languages,” in: Sims-Williams, Nicholas (ed.), *Indo-Iranian Languages and Peoples* (Proceedings of the British Academy, 116): 103–117
- Èdel'man, D. I. 2011–, *Ètimologičeskij slovar' iranskix jazykov*, 4–5, Vostočnaja Literatura, Moscow
- Emmerick, Ronald 1992, “Iranian,” in: Gvozdanović, Jadranka (ed.), *Indo-European Numerals*, Mouton de Gruyter, Berlin: 289–345
- Geiger, Wilhelm & Kuhn, Ernst (ed.) 1895–1903, *Grundriss der iranischen Philologie*, I, Trübner, Strassburg
- Geldner, Karl F. 1886–1896, *Avesta. The Sacred Books of the Parsis*, I–III, Kohlhammer, Stuttgart
- Goto, Toshifumi 2013, *Old Indo-Aryan Morphology and its Indo-Iranian Background*, Verlag der Österreichischen Akademie der Wissenschaften, Vienna
- Henning, W.B. 1958, “Mitteliranisch,” in: Spuler (ed.): 20–130
- Hintze, Almut 2007, *A Zoroastrian Liturgy. The Worship in Seven Chapters* (Yasna 35–41) (Iranica, 12), Harrassowitz, Wiesbaden
- Hoffmann, Karl 1958, “Altiranisch,” in: Spuler (ed.): 1–19. [Reprinted in Hoffmann 1975: 58–76]
- Hoffmann, Karl 1975, 1976, 1992, *Aufsätze zur Indoiranistik*, I–III, Reichert, Wiesbaden
- Hoffmann, Karl 1987, “Avestan language,” in: Yarshater (ed.), 1982–, III/1: 47–62. [Reprinted in Hoffmann 1992: 864–879]

- Hoffmann, Karl & Forssman, Bernhard 2004, *Avestische Laut- und Flexionslehre*, 2nd ed., Institut für Sprachwissenschaft, Innsbruck
- Hoffmann, Karl & Narten, Johanna 1989, *Der Sasanidische Archetypus. Untersuchungen zu Schreibung und Lautgestalt des Avestischen*, Reichert, Wiesbaden
- Humbach, Helmut 1991, *The Gāthās of Zarathushtra and the other Old Avestan Texts*, I–II, Winter, Heidelberg
- Humbach, Helmut & Faiss, Klaus 2010, *Zarathushtra and his Antagonists. A Sociolinguistic Study with English and German Translations of his Gāthās*, Reichert, Wiesbaden
- Inslar, Stanley 1975, *The Gāthās of Zarathustra* (Acta Iranica, 8), Brill, Leiden
- Jackson, A. V. Williams 1892, *An Avesta Grammar in Comparison with Old Indian, Part I: Phonology, Inflection, Word-formation, with an Introduction on the Avesta*, Kohlhammer, Stuttgart
- Kellens, Jean 1984, *Le verbe avestique*, Reichert, Wiesbaden
- Kellens, Jean 1995, *Liste du verbe avestique*, Reichert, Wiesbaden
- Kellens, Jean & Pirart, Eric 1988, 1990, 1991, *Les textes vieil-avestiques*, I–III, Reichert, Wiesbaden
- Kent, Roland G. 1953, *Old Persian. Grammar, Texts, Lexicon*, 2nd ed., American Oriental Society, New Haven
- Martínez, Javier & de Vaan, Michiel 2014, *Introduction to Avestan*, Brill, Leiden
- Mayrhofer, Manfred 1978, *Supplement zur Sammlung der altpersischen Inschriften*, Verlag der Österreichischen Akademie der Wissenschaften, Vienna
- Mayrhofer, Manfred 1983, “Lassen sich Vorstufen des Uriranischen nachweisen?,” *Anzeiger der phil.-hist. Klasse der Österreichischen Akademie der Wissenschaften* 120: 249–255
- Morgenstierne, Georg 1973, *Irano-Dardica*, Reichert, Wiesbaden
- Morgenstierne, Georg 1973a, “Traces of Indo-European accentuation in Pashto?,” *Norsk Tidsskrift for Sprogvidenskap* 27: 61–65
- Narten, Johanna 1968, “Zum ‘proterodynamischen’ Wurzelpresens,” in: J. C. Heesterman, G. H. Schokker & V. I. Subramoniam (eds.), *Pratidānam. Indian, Iranian and Indo-European Studies presented to F. B. J. Kuiper*, Mouton, The Hague: 9–19
- Narten, Johanna 1986, *Der Yasna Haptaṅhāiti*, Reichert, Wiesbaden
- Rastorgueva, V. S. & Édel’man, D. I. 2000–2007, *Ètimologičeskij slovar’ iranskij jazykov*, 1–3, Vostočnaja Literatura, Moscow
- Reichert, Hans 1909, *Avestisches Elementarbuch*, Winter, Heidelberg
- Schmitt, Rüdiger 1987, “Aryans”, in: Yarshater (ed.), 1982–, II/7: 684–687
- Schmitt, Rüdiger (ed.) 1989, *Compendium Linguarum Iranicarum*, Reichert, Wiesbaden
- Schmitt, Rüdiger 1991, *The Bisitun Inscriptions of Darius the Great. Old Persian Text*, School of Oriental and African Studies, London
- Schmitt, Rüdiger 2000, *The Old Persian Inscriptions of Naqsh-e Rostam and Persepolis*, School of Oriental and African Studies, London
- Schmitt, Rüdiger 2014, *Wörterbuch der altpersischen Königsinschriften*, Reichert, Wiesbaden
- Sims-Williams, Nicholas 2007, “The Sogdian potentialis,” in: Maria Macuch, Mauro Maggi & Werner Sundermann (eds.), *Iranian Languages and Texts from Iran and Turan. Ronald E. Emmerick Memorial Volume* (Iranica, 13), Harrassowitz, Wiesbaden: 377–386
- Skjærvø, Prods Oktor 2009, “Old Iranian,” in: Windfuhr (ed.) 2009: 43–195

- Spuler, Bertold (ed.) 1958, *Handbuch der Orientalistik, 1. Abt., Bd. IV: Iranistik, 1. Abschnitt: Linguistik*, Brill, Leiden-Köln
- West, Martin L. 2011, *Old Avestan Syntax and Stylistics. With an Edition of the Texts* (Abhandlungen der Akademie der Wissenschaften zu Göttingen, Neue Folge, 13), de Gruyter, Berlin
- Windfuhr, Gernot (ed.) 2009, *The Iranian Languages* (Routledge Language Family Series), Routledge, London
- Yarshater, Ehsan (ed.) 1982, *Encyclopaedia Iranica*, I–IV, Routledge and Kegan Paul, London; V–VIII, Mazda Publishers, Costa Mesa; IX–XI, Bibliotheca Persica Press, New York; XII–, Encyclopaedia Iranica Foundation, New York

# GREEK

*Rupert Thompson*

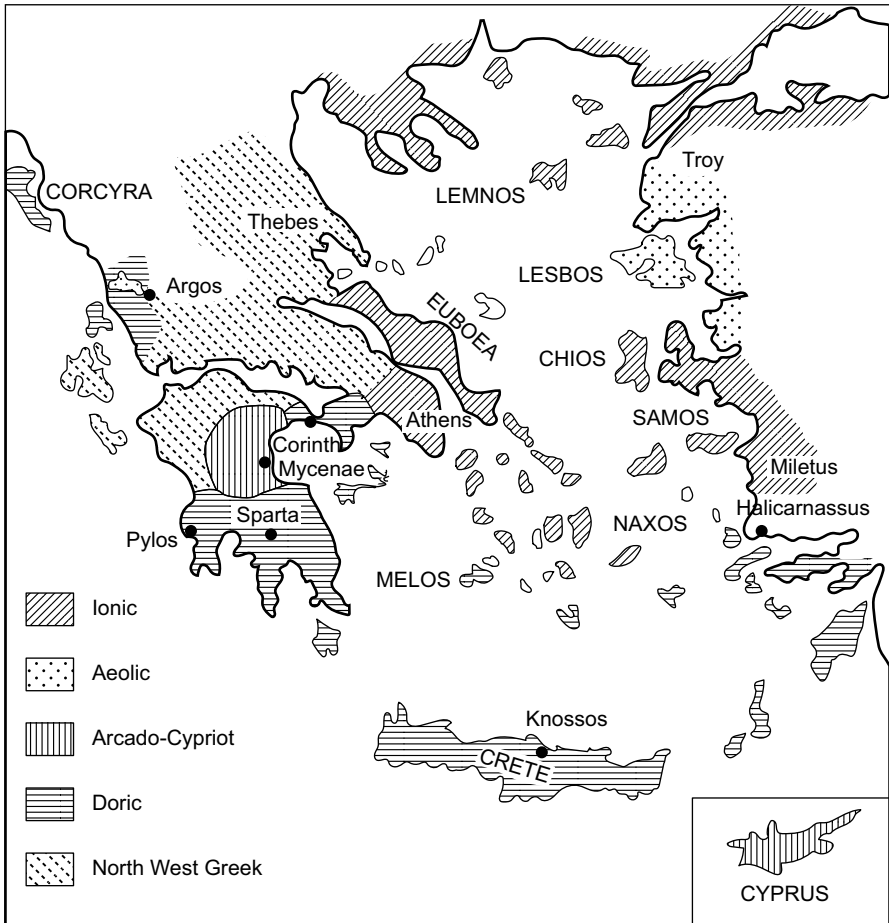
## INTRODUCTION

Greek is today the official language of Greece and of the Republic of Cyprus, and is spoken by sizeable diaspora communities across the globe. Varieties of Greek are spoken by minorities in southern Italy (Griko) and on the coast of the Black Sea (Pontic). Until the early part of the twentieth century there were large numbers of Greek speakers in Turkey; dialects such as Cappadocian are now all but extinct. All of the modern dialects of Greek, with the possible exception of Tsakonian (spoken by a declining minority in villages on the slopes of Mount Parnon), which may be a survival of the ancient Laconian dialect, descend from the variety known as the Koine (Gr. κοινή διάλεκτος, ‘common dialect’) which itself developed from the “internationalised” version of the dialect of Athens of the fifth century BC, “Great Attic”, and was adopted by the Macedonian kings as the official court language. From an IE perspective Greek is a member of the “core” group of languages left behind by the split of Anatolian and Tocharian. It is generally now considered that it developed in situ in the Balkan Peninsula in the second millennium BC. A number of words, including some place names, terms for flora and fauna, and cultural items (Κόρινθος ‘Corinth’, σῦκον ‘fig’, ἀσάμινθος ‘bath-tub’) attest substrate influence of non-Greek (even non-IE) languages. Some, with problematic phonological details (e.g. ἵππος ‘horse’ < PIE \*ekwos, with unexplained aspiration and vowel quality), suggest contact with otherwise unknown IE varieties.

Before the prestige of the Koine caused their ultimate demise in the Hellenistic period there had existed in the Greek-speaking world a large variety of regional dialects. None had the status of a standard language; the Greeks were clear enough that they all spoke varieties of Greek – famously according to Herodotus 8.144 τὸ Ἑλληνικόν ‘Greekness’ is ὁμόγλωσσον ‘of one language’ – but “Greek” in this sense was no more than an abstraction (Morpurgo Davies 1987). Certain dialects, often with extreme local features watered down or altogether removed, did become the conventional medium for certain genres of literature. For example, no matter what the native dialect of the poet, choral lyric – even the lyric passages of Attic tragedy – was written in a literary version of Doric (although in tragedy this largely amounts to the replacement of Att. η by Dor.  $\bar{a}$  and the use of the Dor. first-declension gen. pl. in - $\bar{a}\nu$  rather than Att. - $\bar{\omega}\nu$ ). Epic poetry was composed in the largely Ionic-based Homeric *Kunstsprache*.

The dialects have traditionally been classified into two main families, designated East Greek and West Greek according to their distribution in the first millennium. East Greek comprises Attic-Ionic and the more conservative Arcado-Cypriot; West Greek consists of (Peloponnesian) Doric and North-West Greek. A third group, Aeolic, made up of Thesalian and Boeotian on the Greek mainland, and Lesbian, spoken on the island of Lesbos and the facing northern part of the coast of Asia Minor, is difficult to fit into this binary classification. García Ramón (1975) argues that these dialects arose from contact between East and West Greek varieties; recent scholarship (following Parker 2008) calls

into question the very existence of an Aeolic group and suggests that the three dialects in question may be unrelated archaic varieties. The dialect of Pamphylia is similarly hard to place (see Brixhe 1976). On the question of whether or not ancient Macedonian was a Greek dialect – a matter of no less heated debate in the ancient world than it is today – see Mendéz Dosuna 2012.



**MAP 5.1** THE GREEK DIALECTS

Source: Adapted from: L. R. Palmer, *The Greek Dialects*, London: Faber and Faber, 1980

From the second half of the eighth century BC (the earliest examples are the “Dipylon pitcher” from the Kerameikos cemetery in Athens, dated to ca 740 BC, and the “cup of Nestor” from Pithekoussai on the Italian island of Ischia, dated to 750–700 BC) inscriptional material is written in a number of related local (“epichoric”) alphabets, all deriving ultimately from an adaptation of the Phoenician abjad. Gradually they came to be replaced by a standardised alphabet of East Ionic origin (Table 5.1). Modern conventional orthography mixes upper- and lower-case graphs and employs diacritics of Alexandrine

and Byzantine origin for the accent, for aspiration (or the lack of it) on vowels and *r*, and for the second member *i* of long diphthongs. For details of the epichoric alphabets see Jeffery 1990. In addition to the epigraphic record the contemporary written material is augmented by a large number of papyri from Egypt dating to the Hellenistic and Roman periods. These include a variety of personal and official documents which provide invaluable evidence of the linguistic behaviours of a broad spectrum of social strata. They also include early copies of some literary texts. The literary evidence has otherwise been mediated by a long tradition of manuscript copyists.

**TABLE 5.1 THE GREEK (IONIC) ALPHABET.** Phonetic values are those of Attic ca 400 BC. The equivalent spellings in the Old Attic alphabet, used in Athens before 403 BC, are also shown.

Grapheme	Name	Phonetic value	Old Attic equivalent	Notes
A, α	alpha	[a, a:]	A	
B, β	beta	[b]	B	
Γ, γ	gamma	[g] / [ŋ]	Γ	(1)
Δ, δ	delta	[d]	Δ	
E, ε	epsilon	[e]	E	
F, Ϝ	digamma (wau)	[w]	—	(2)
Z, ζ	zeta	[zd]	Z	
H, η	eta (heta)	[ɛ:]	E (H=[h])	(3)
Θ, θ/ϑ	theta	[tʰ]	Θ	
I, ι	iota	[i, i:]	I	
K, κ	kappa	[k]	K	
Λ, λ	lambda	[l]	Λ	
M, μ	mu	[m]	M	
N, ν	nu	[n]	N	
Ξ, ξ	xi	[ks]	XΣ	
O, ο	omicron	[o]	O	
Π, π	pi	[p]	Π	
P, ϱ	rho	[r, rʰ]	P	(4)
Σ, σ/ς	sigma	[s]	Σ	(5)
T, τ	tau	[t]	T	
Υ, υ	upsilon	[ü, ü:]	Y	(6)
Φ, φ	phi	[pʰ]	Φ	
X, χ	chi	[kʰ]	X	
Ψ, ψ	psi	[ps]	ΦΣ	
Ω, ω	omega	[ɔ:]	O	
ει		[e:]	EI = [ei], E = [e:]	(7)
ου		[u:]	OY = [ou], O = [o:]	(7)
α		[a:(i)]	AI	(8)
η		[ɛ:(i)]	EI	(8)
φ		[ɔ:(i)]	OI	(8)

**Notes to Table 5.1:**

- (1) Before κ, χ, μ or another γ the letter γ has the articulation of a velar nasal [ŋ].
- (2) The sound [w] was lost prehistorically in Attic-Ionic, and consequently the grapheme <F> was not used in those dialects. The name “digamma” comes from its shape, which resembles two capital gammas superimposed.

- (3) In most varieties of the Greek alphabet the grapheme <H> has the value [h]. East Ionic, lacking the phoneme *h*, redeployed the redundant sign as a grapheme for the long vowel which became  $\bar{\epsilon}$ . (See below under “Phonology”.) Current practice, originating in Alexandrian notation, is to print the *rough breathing* <’> above initial vowels (and *r*) which have aspiration and the *smooth breathing* <˘> above those which lack it.
- (4) Word-initially, and perhaps in the second member of a geminate, *r* is voiceless and aspirated, [rʰ]. Elsewhere it is voiced and unaspirated [r]. Word-initial rho is thus traditionally written with the rough breathing,  $\rho$ .
- (5) The lower-case form  $\sigma$  is used medially,  $\varsigma$  word-finally. The ancient Greeks themselves used only upper-case letter forms, and so made no such distinction. Consequently, some scholars prefer to use “lunate” sigmas C, c, with the same lower-case form used in all environments.
- (6) In most dialects, and originally in Attic, the value of  $\upsilon$  is [u, u:]. See below under “Phonology”.
- (7) The sequences  $\epsilon\iota$  and  $\omicron\upsilon$  are known as “spurious diphthongs”. Originally they represented diphthongs [ei] and [ou], but sometime before 403 BC in Attic they had monophthongised to close-mid vowels [e:, o:] (transcribed here  $\bar{\epsilon}$ ,  $\bar{\omicron}$ ), the latter subsequently raising to [u:].
- (8) These represent original “long diphthongs”  $\bar{a}i$ ,  $\bar{\epsilon}i$ ,  $\bar{o}i$  – see Allen 1987a: 84–88 for the likely phonetic reality of these sounds – which were in inscriptions written AI, EI (later HI), OI (later  $\Omega$ I), but which by ca 200 BC lost the second element to merge with  $\bar{a}$ ,  $\bar{\epsilon}$ ,  $\bar{o}$  and were then written A, H,  $\Omega$ . The convention of indicating this lost *i* with an *iota subscriptum* is of Byzantine origin and is generally followed today, except in the case of capitals, e.g.  $\text{Αἰδης} = \text{Hāidēs}$  ‘Hades’: were it *Haidēs* the accent and breathing would be printed over the *iota*: Αἶδης. Some scholars prefer to print such *iotas* adscript,  $\bar{a}i$ ,  $\eta i$ ,  $\omega i$ .

Before 1952 the earliest Greek material known to us was provided by the two epic poems attributed in antiquity to Homer, the *Iliad* and the *Odyssey*. The language of these poems is broadly Ionic, but has an admixture of archaic, Aeolic and wholly artificial forms. Thanks to the work of Milman Parry (Parry 1928a, 1928b; see also Parry et al. 1971) it is known that this Homeric *Kunstsprache* owes its formation to a long tradition of oral poetry, dating back to the Bronze Age and beyond. The poems are composed in strict dactylic hexameters, and to assist the process of oral composition the poets had evolved a complex system of *formulae*, “building-blocks” of convenient metrical shape, e.g.  $\text{πολύτλας δῖος Ὀδυσσεύς}$  ‘much-suffering godlike Odysseus’ or  $\text{πολύμητις Ὀδυσσεύς}$  ‘much-cunning Odysseus’, which fill in the shapes, so conveniently completing the line from the weak third-foot caesura and strong fourth-foot caesura respectively. Where such *formulae* contained material which was no longer linguistically current the Ionic poets replaced it with contemporary Ionic forms if they had the same metrical shape; otherwise with suitable contemporary forms from a parallel Aeolic tradition; and failing that they retained the archaic material (so Horrocks 1987, 1997; but the idea that the tradition went through a distinct Aeolic phase before passing entirely into Ionian hands is not totally laid to rest. For discussion see the contributions by Janko and Jones in Andersen and Haug 2012). The result is an amalgam from various time-depths, containing, for example, contemporary AI *o*-stem gen. sg. in  $-\omicron\upsilon$  alongside its earlier form,  $-\omicron\iota\omicron$ ; and some features, such as “*tmesis*” – the separation of a preverb from its verb – and the scansion of  $\text{ἀνδρότητα}$  ‘manhood’ at *Iliad* 16.857

as if \*anḡtāta, must pre-date the Bronze Age. The poems took their present form probably in the eighth or seventh century BC, and it is doubtful whether there was a single poet responsible for each poem, let alone both; but the name “Homer” is a convenient shorthand for the poetic tradition.

Excavating in Knossos on Crete in 1900, Sir Arthur Evans discovered clay tablets written in an unknown script, today called Linear B, dating to ca 1400 BC. Similar documents of a slightly later date were unearthed in subsequent excavations at Pylos in Messenia, at Mycenae in the Argolid and at Thebes in Boeotia, and continue to be found to the present day, the most recent discoveries coming from Ayios Vasilios in Lakonia. A collaboration in the 1950s between Michael Ventris and John Chadwick, building on earlier work by the American scholar Alice Kober, revealed that these documents were the economic records of Bronze Age palatial elites written in a very early form of Greek which today we call Mycenaean. Despite its early date and obvious conservatism in some areas (e.g. retention of PIE labiovelars and \*w, preservation of instrumental morphology) Mycenaean has undergone the assibilation of Proto-Greek \*ti to *si* which characterises the East Greek dialect family, indicating that the split into East and West Greek had taken place already in the second millennium.

Linear B evidently derives from an earlier script, Linear A, used to write an unknown language of Crete (and which itself superseded a more elaborate writing system called Cretan Hieroglyphic). The script is not especially well adapted to writing Greek. In addition to a large number of ideograms it contains some 88 syllabic signs, each of which denotes a simple vowel, a sequence of consonant + vowel or, in a small number of cases, more complex sequences such as *rai* or *nwa*. It does not normally distinguish between voiceless, voiced and aspirated stops (except that *t* and *tʰ* are distinguished from *d*, and there is an optional sign representing *pʰu*), nor between *r* and *l*, and vowel length is not marked. Consonant clusters cannot be written: the scribes either omit consonants (“partial spelling”, e.g. *pe-ma* = *sperma* ‘grain’) or insert “dummy vowels” (“plene spelling”, e.g. *wa-na-ka-te* = *wanaktei* ‘for the king’). Diphthongs are not always marked, and word-final consonants almost never are.

Another syllabic script derived ultimately from Linear A was used to write the ancient dialect of Cyprus. The earliest example may be the spit inscribed with the gen. of the man’s name *Opʰeltau* dating from the tenth century BC, although it has recently been suggested that this may be the earlier “Cypro-Minoan” script (used to write one or more non-Greek languages). The Cypriot script differs from Linear B in lacking signs for voiced dentals, and in its strategies for writing consonant clusters.

## PHONOLOGY

The consonantism of Greek is fairly conservative. Being a *centum*-type language, the PIE velars and palatals have merged as velars. The major development of the consonant system which characterises Proto-Greek is the loss of voicing on the PIE voiced aspirates, PIE \*bʰ, \*dʰ, \*gʰ becoming φ, θ, χ = [pʰ, tʰ, kʰ]. The voiceless aspirated stop articulation of these phonemes is preserved into the Roman period: the earliest evidence for a pronunciation [f] for φ and [θ] for θ is from the first century AD. (This is true at least for Attic and its descendant, the Koine. In Laconian θ may have become a fricative by the end of the fifth century. In the *Lysistrata*, Aristophanes has the Spartans use forms such as σιός for Attic θεός, perhaps an attempt to write [θ].) Similarly, β, δ, γ retained their articulation as voiced stops throughout the ancient period; it is



not clear when they became the fricatives [β, δ, γ] of Modern Greek, but it may be as late as the ninth century AD.

### Labiovelars

The PIE labiovelars are originally retained, except when adjacent to *ǔ* or *w* when they lose their labialisation and merge with the plain velars. Mycenaean spells all labiovelars with a series of signs transcribed *qa, qe, qi, qo*, e.g. *a-pi-qa-ro* = *amp<sup>h</sup>ik<sup>w</sup>olos* (Att. ἀμφίπολος) ‘attendant’ while *qa-si-re-u* = *g<sup>w</sup>asileus* (Att. βασιλεύς, a local official in Mycenaean, ‘king’ in later Greek). In *qo-u-ko-ro* = *g<sup>w</sup>oukolos* (Att. βουκόλος) ‘cowherd’ the first labiovelar is retained, and the second has become velar because of the preceding *u*.

In the dialects of the first millennium the remaining labiovelars merged with various other consonantal phonemes. In Attic-Ionic and West Greek the development is twofold: (i) before *e* (and in the case of *\*k<sup>w</sup>* also before *i*) the outcome is dental, *\*k<sup>w</sup>* > *t*, *\*g<sup>w</sup>* > *d*, PIE *\*g<sup>wh</sup>* > Proto-Greek *\*k<sup>wh</sup>* > *t<sup>h</sup>*; (ii) elsewhere the outcome is labial, *\*k<sup>w</sup>* > *p*, *\*g<sup>w</sup>* > *b*, PIE *\*g<sup>wh</sup>* > Proto-Greek *\*k<sup>wh</sup>* > *p<sup>h</sup>*. This is the major source of the *b* phoneme in Greek. Examples are given in Table 5.2. Other dialects behave differently. In Aeolic, for example, the outcome is labial across the board (e.g. thus Lesb. πέμπτε = Att. πέντε ‘five’), but even in Aeolic the pronominal stem *\*k<sup>w</sup>i-* gives τίς etc.

TABLE 5.2 TREATMENT OF LABIOVELARS IN ATTIC-IONIC AND WEST GREEK

	<i>*k<sup>w</sup></i>	<i>*g<sup>w</sup></i>	<i>*g<sup>wh</sup></i> > Proto-Greek <i>*k<sup>wh</sup></i>
before <i>e</i>	<i>*penk<sup>w</sup>e</i> > πέντε ‘five’	<i>*g<sup>w</sup>elb<sup>h</sup>u-</i> > δελφύς ‘womb’	<i>*g<sup>wh</sup>en-</i> > θείνω ‘kill’
before <i>i</i>	<i>*k<sup>w</sup>is</i> > τίς ‘who?’	<i>*g<sup>w</sup>ih<sub>3</sub>-</i> > βίος ‘life’	<i>*h<sub>3</sub>eg<sup>wh</sup>i-</i> > ὄφις ‘snake’
elsewhere	<i>*penk<sup>w</sup>tos</i> > πεμπτός ‘fifth’	<i>*g<sup>w</sup>ῆ-</i> > βαίνω ‘go’	<i>*g<sup>wh</sup>on-</i> > φόνος ‘murder’

### Laryngeals

The behaviour of the PIE laryngeals is similar in Greek to in the other daughter languages, with two important exceptions. First, and unique to Greek, is the so-called triple-reflex of laryngeals in positions where they vocalise. In all other IE languages all three laryngeals give the same reflex, *i* in Indo-Iranian, *a* elsewhere; in Greek, however, *\*h<sub>1</sub>* vocalises to *e*, *\*h<sub>2</sub>* to *a*, *\*h<sub>3</sub>* to *o*. Thus, *\*d<sup>h</sup>h<sub>1</sub>tos* > θετός ‘set’, *\*sth<sub>2</sub>tos* > στατός ‘placed’, *\*dh<sub>3</sub>tos* > δοτός ‘granted’. Second, word-initial *\*H* before a resonant other than *\*y* also vocalises rather than being lost as in other languages (other than Armenian): *\*h<sub>1</sub>rud<sup>h</sup>-* > ἐρυθρός ‘red’, *\*h<sub>2</sub>ner-* > ἀνήρ ‘man’, *\*h<sub>3</sub>nom<sup>h</sup>-* > ὄνομα ‘name’. This is the origin of at least some *prothetic vowels* in Greek (compare Lat. *ruber*, *Nero* and *nōmen* for the equivalent forms without). The triple reflex is also seen in the outcome of *\*RH*, as ρη, ρᾶ, ρω etc.

### Semivowels

PIE *\*w* was inherited into Proto-Greek, is preserved in Mycenaean, but was lost at various stages in the later dialects. Where retained, it was written with the letter Ϝ. In Attic-Ionic its disappearance was prehistoric. It was originally present in the Homeric tradition, where at the convenience of the poet it is either observed – to block hiatus (e.g. μελιγδέα (Ϝ)οῖνον ‘honey-sweet wine’) or make a preceding vowel long by position (e.g. εἴπας (Ϝ)ἔπος) – or ignored (e.g. μελιγδέος οἴνου).

Medially, PIE \*y was lost via [h]. Word-initially it shows a double treatment. In some roots it gives Greek ζ- (originally with the value [dʒ]), in others [h]. The outcome is consistent across the dialects. Those roots which give later Greek ζ- are in Mycenaean spelt with signs of the z-series (e.g. PIE \*yewg- > Myc. *ze-u-ke-si* = *dzeuges(s)i* ‘pair (dat. pl.)’, Att. ζεύγεσι); those which give [h] are spelt with signs of the j-series or simple vowels (e.g. from the \*yo- pronominal stem adverbs *jo-* = *yō* and *o-* = *hō* ‘how’). This probably shows that the change \*y > h was in progress at the time the Linear B documents were written. The alternation between h- and ζ-roots has not been satisfactorily explained, although as \*Hy- does not give a prothetic vowel, word-initial laryngeals have sometimes been supposed to underlie one development or the other.

Later Greek ζ has two other sources: (i) the palatalisation of \*dy and \*gy, e.g. \*ped-yos > πεζός ‘on foot’, \*meg-yos- > Ion. μέζων, comp. of μέγας ‘big’; and (ii) the combination of s+d, e.g. Ἀθήναζε < Ἀθήνασ+δε ‘to Athens’. The former type is spelt in Mycenaean with z-series signs (e.g. *me-zo-e* = *medzohes* ‘bigger’), the latter with signs of the d-series (e.g. *te-qa-de* = *T<sup>h</sup>ēg<sup>w</sup>ans-de* ‘to Thebes’). This indicates that the original outcome of \*dy and \*gy metathesized to *zd* and thus merged with *s+d* in the post-Mycenaean period.

Voiceless \*t<sup>(h)</sup>y and \*k<sup>(h)</sup>y also underwent palatalisation. There are two distinct phases. First, those sequences of \*ty and \*thy where no morpheme boundary intervened palatalised to [ts], which gives -*tt-* in Boeotian, -*ss-* elsewhere, simplifying to -*s-* in Attic-Ionic, e.g. \*yotyos > ΑΙ ὄσος ‘how big’ vs. Lesb., Thess., Dor. (h)οσσος; PIE \*med<sup>h</sup>ynos > Proto-Greek \*met<sup>h</sup>ynos > ΑΙ μέσος ‘middle’, Boe. μεττος, elsewhere μεσσος. Where a morpheme boundary intervenes the change is delayed. For most dialects the outcome is the same as homomorphemic \*t<sup>(h)</sup>y (viz. Boe. -*tt-*, general -*ss-*), but Attic this time follows the Boeotian treatment (e.g. \*melit-ya > Att. μέλιττα ‘honey bee’) and in Ionic the -*ss-* does not simplify (e.g. μέλισσα). The outcomes of \*k<sup>(h)</sup>y and \*tw are the same: Proto-Greek \*p<sup>h</sup>ulak-yō ‘guard’ > Att. and Boe. φυλάττω vs. φυλάσσω elsewhere; and Proto-Greek \*k<sup>w</sup>etw<sup>r</sup>- ‘four’ > Att. τέτταρες, Boe. πετταρες vs \*k<sup>w</sup>etwer- > Ion. τέσσερες. Interestingly, words of non-Greek origin show the same distribution of -*tt-* vs -*σσ-*, e.g. Att. and Boe. θάλαττα vs. general θάλασσα ‘sea’.

### The PIE fricative

Between vowels, and before a vowel at word-beginning, PIE \*s > h, and medially is then lost. The lenition of \*s > h is pre-Mycenaean, the subsequent loss post-Mycenaean: from Proto-Greek \*p<sup>h</sup>arwesa ‘cloths’, Mycenaean has *pa-we-a<sub>2</sub>* = *p<sup>h</sup>arweha*. This has interesting implications for Grassmann’s Law, by which the first in a sequence of two aspirates loses aspiration (e.g. \*sek<sup>h</sup>- > Proto-Greek \*hek<sup>h</sup>ō > Att. pres. ἔχω, vs. fut. \*hek<sup>h</sup>sō > \*heksō > ἔξω). Since \*d<sup>h</sup>esos > Proto-Greek \*t<sup>h</sup>ehos > θεός ‘god’ we can perhaps infer that, although it is common to all dialects of the first millennium, Grassmann’s Law operates after the loss of medial -*h-*, i.e. in the post-Mycenaean period. Otherwise, the classical form would be \*τέός.

### The vowel system

Proto-Greek inherited from PIE a simple vowel system with five short vocalic phonemes, *a*, *e*, *i*, *o*, *u*, each with a corresponding long equivalent. There were also diphthongs *ai*, *ei*, *oi*, *au*, *eu*, *ou*, and long diphthongs *āi*, *ēi*, *ōi*, *āu*, *ēu*, *ōu*. Early in the history of the

language, perhaps at the Proto-Greek stage, certain medial clusters involving a liquid or nasal and *s* underwent loss of the *s* and gemination of the resonant. Mycenaean attests this stage; thus, e.g., from Proto-Greek \*agersantes, the aor. ptcp. m. nom. pl. of the verb ἀγείρω ‘collect’, Mycenaean has *a-ke-ra<sub>2</sub>-te = agerrantes*. The geminate is preserved in Lesbian and Thessalian, but the other dialects further simplify the cluster by degemination with compensatory lengthening of a preceding short vowel. In the case of *a*, *i* and *u* the outcome is identical to the existing long vowels; this is also true of *e* and *o* in Boeotian, Arcadian (and one supposes Cypriot, although the writing system obscures the details) and some Doric dialects labelled by Ahrens (1843: 403–414) as *Doris seuerior* (‘severer Doric’); in other dialects, however, including Attic, Ionic and those Doric varieties labelled as *Doris mitior* (‘milder Doric’), *e* and *o* lengthened not to inherited  $\bar{e}$  and  $\bar{o}$  (which were [ɛː, ɔː]) but to new close-mid vowels  $\bar{e}$  and  $\bar{o}$  (phonetically [eː, oː]). Allen (1959, 1987b) suggests that in these varieties the short-vowel system was skewed towards the top of the vocalic space (Figure 5.1), so that short *e*, *o* were of noticeably closer quality than inherited long  $\bar{e}$ ,  $\bar{o}$ . He cites in support the fact that in these dialects PIE \* $\text{r}_1$  and \* $\text{r}_2$  give *ar~ra* and *al~la*: the original outcome is likely to have been [ar] etc., and the top-skewing would mean that [ə] was sufficiently close to *a* as to merge with it. The result of this is shown in Figure 5.2. The new  $\bar{e}$  and  $\bar{o}$  phonemes were spelt in Old Attic orthography as <E, O>. After 403 BC and in standard modern orthography they are spelt *ei* and *ou* for reasons which are explained below. Examples of the treatment of various clusters are:

Proto-Greek \*k<sup>h</sup>eslioi > Ion. χεῖλιοι, ‘severe’ Dor. χηλιοι, Lesb. and Thess. χέλλιοι ‘thousand’

Proto-Greek \*asmes > AI ἡμεῖς (< \*ḗμεις), Dor. ἄμες, Lesb. ἄμμες ‘we’

Proto-Greek \*esmi > AI εἰμί, ‘severe’ Dor. ἦμι, Lesb. ἔμμι ‘I am’

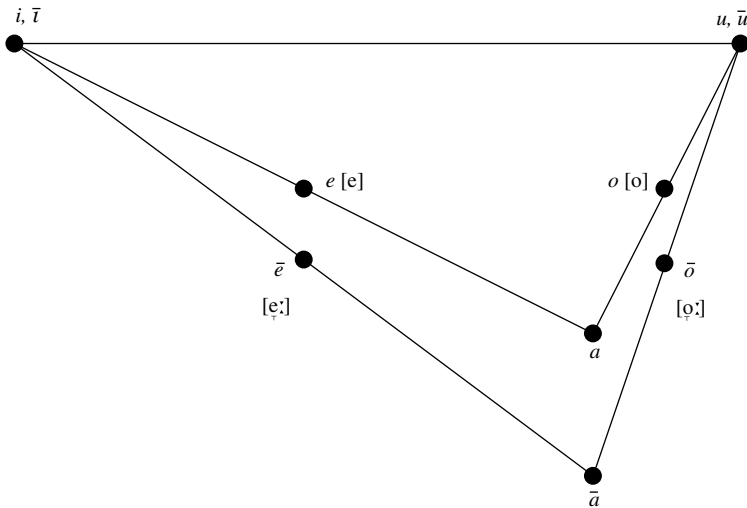


FIGURE 5.1 INHERITED VOWEL SYSTEM OF PROTO-ATTIC-IONIC

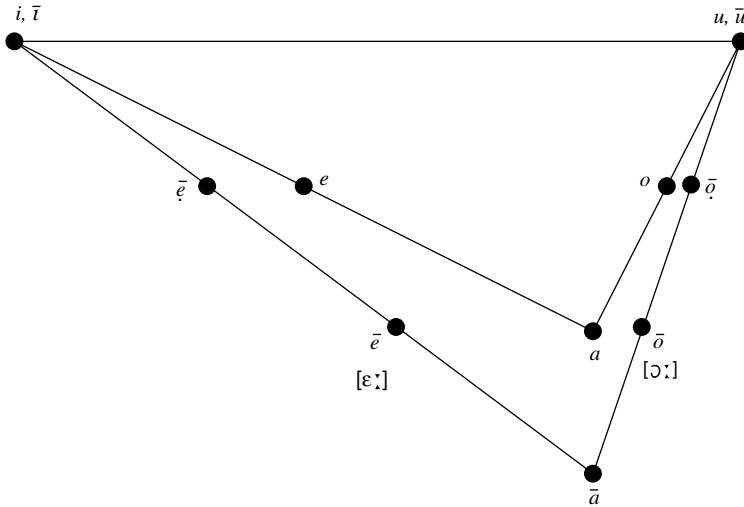


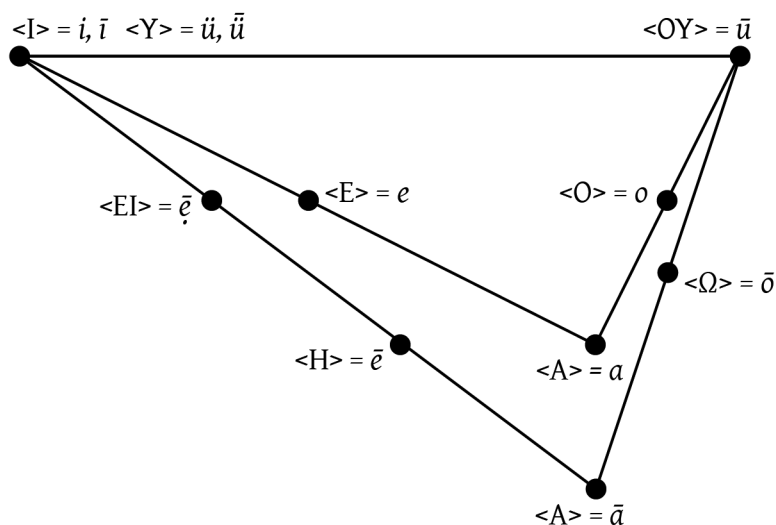
FIGURE 5.2 VOWEL SYSTEM OF PROTO-ATTIC-IONIC AFTER COMPENSATORY LENGTHENING

The resulting system seems to have suffered from overcrowding on the back axis, which is substantially shorter than the front. In Proto-Attic-Ionic this had implications for both the top and bottom of the back axis. First, the  $\bar{a}$  phoneme underwent fronting to become [æ:]. A second wave of compensatory lengthenings, this time involving word-final *ns* clusters and those with secondary *s*, added to the inventory of  $\bar{e}$  and  $\bar{o}$  and reintroduced a new  $\bar{a}$ . It is at this stage that Eastern Ionic redeployed the redundant grapheme <H> to represent the shifted  $\bar{a}$  phoneme: a seventh-century BC inscription from Naxos spells the word ‘sister’, phonetically [kasigné:tæ:], as ΚΑΣΙΓΝΕΤΗ, using <E> for inherited  $\bar{e}$  (and also for *e* and  $\bar{e}$ ) and <H> for  $\bar{a}$  <  $\bar{a}$ . <A> is used for inherited *a* and the new  $\bar{a}$  from the second wave of compensatory lengthening. Eventually, in Ionic,  $\bar{a}$  merged completely with  $\bar{e}$ , taking its grapheme with it, so the standard spelling of the ‘sister’ word became κασιγνήτη, representing [kasigné:tɛ:]. In Attic it split, merging with the new  $\bar{a}$  after *e*, *i* or *r*, and with  $\bar{e}$  elsewhere.

The effects of overcrowding were now felt again, this time causing a shift in the top of the back axis. Inherited  $\bar{u}$  (spelt <Y>) was pushed over onto the front axis, becoming  $\bar{u}$  (a high front vowel with lip-rounding), and for reasons of symmetry the short *u* followed it to become  $\bar{u}$ . The original value of <Y> is seen in the onomatopoeic verb  $\mu\kappa\acute{\alpha}\omicron\mu\alpha\iota$  ‘moo’, and in the spellings <AY, EY> of the diphthongs *au*, *eu*. The close-mid  $\bar{o}$  then moved up to occupy the space so vacated at the top of the back axis and became [u:]. This is its value in Classical Attic. As a consequence there is no longer a corresponding short *u* phoneme. These developments explain why the Boeotians, when they adopted the Attic alphabet in the fourth century, considered <Y> unsuitable for writing their own phonemes which had retained the values [u, u:] and instead used the digraph <OY>. Occasionally, too, spellings <AO, EO> are found for *au*, *eu*.

In the Old Attic orthography <E> was used not only for  $\bar{e}$  but also  $\bar{e}$  and  $\bar{e}$ , and <O> likewise for  $\bar{o}$ ,  $\bar{o}$  and  $\bar{o}$  (>  $\bar{u}$ ). In 403 BC, under the archonship of Euclides, Athens adopted the Ionic alphabet, which used <H> and <Ω> for  $\bar{e}$  and  $\bar{o}$ . At the same time

<EI> and <OY> were employed for  $\bar{e}$  and  $\bar{u}$  <  $\bar{o}$ . These spellings originally represented the diphthongs [ei] and [ou], which had by the fifth century BC monophthongised and merged with  $\bar{e}$  and  $\bar{o}$  (>  $\bar{u}$ ), and it is unsurprising that they should then be used for the simple vowels from whatever source. They are traditionally called “spurious diphthongs” but are properly speaking digraph spellings of monophthongs. The vowel system of Classical Attic, with the post-Eucleidean spellings, is shown in Figure 5.3.



**FIGURE 5.3** VOWEL SYSTEM OF CLASSICAL ATTIC, SHOWING OFFICIAL SPELLINGS AFTER 403 BC

It is worth saying a few words about the later developments of the vowel system into the Koine and beyond. Around the end of the fourth century BC  $\bar{e}$  raised and merged with  $\bar{i}$ ; one imagines that  $\bar{e}$  itself then raised slightly from [e:] to [ɛ:]. By the end of the first century AD a new wave of monophthongisations caused *oi* to merge with  $\bar{u}$ , and turned *ai* into [ɛ:]; in turn,  $\bar{e}$  raised to [e:] and eventually [i:], merging with  $\bar{i}$ . Between the second and third centuries AD phonemic vowel length was lost, probably in connection with the change from a pitch to a stress accent, and, finally, sometime after the fourth century  $\bar{u}$  (including the outcome of earlier *oi*) lost its lip-rounding and merged with *i*, giving the simple five-vowel system of Modern Greek: *a* <α>, *e* <ε, αι>, *i* <ι, ει, η, ι, υ, οι>, *o* <ο, ω>, *u* <ου>.

### Accent

The details of the Greek accentual system are beyond the scope of the present chapter. The reader is referred to Probert 2003. Only a basic overview is presented here.

From the time of Plato ancient commentators on the accentual system of Greek describe two categories of accent, ὀξύς ‘sharp’ and βαρύς ‘heavy’. Two of the general terms for accent, τόνος and τάσις, both meaning ‘stretching’, are metaphors of the tension in the strings of a musical instrument, and the third, προσῳδία ‘singing-along-with’, also

implies an accent based on pitch rather than stress. The categories ὀξύς and βαρύς are themselves associated with the terms ἐπίτασις ‘stretching’ and ἄνεσις ‘slackening’, and are thus clearly high and low pitch respectively.

The tradition of marking the accent with diacritics originates in Alexandria ca 200 BC, but the present system of notation is Byzantine. The acute marks a rise in pitch followed by a fall in pitch on the next syllable, e.g. ἄνθρωπος with rise on the *α* and fall on the *ω*. Each change of pitch requires one mora (the length of a short vowel). If the accent is borne by a long vowel or diphthong (both of which equal two morae) the pitch may rise on the second mora and then fall on the following syllable, in which case it is still marked acute (e.g. ἀνθρώπου, which may be regarded schematically as *ant<sup>h</sup>roópiuu*), or it may rise on the first mora and fall on the second mora of the same syllable, in which case it is marked with the circumflex (e.g. πῶγμα, schematically *práagma*). When a word which bears an acute on its final syllable (the traditional term is “oxytone”) is followed by another accented word, its accent is written grave (barytonesis, e.g. ἀγαθὸς ἄνθρωπος). It is not clear what this represents, but it is a reasonable assumption that the pitch was required to return to the base level by the word end, and that the grave accent indicates that the expected rise in pitch is suppressed, or at least moderated.

The position of the accent is free within certain constraints, known collectively as the law of limitation: if the final syllable is light (i.e. contains a short vowel followed by at most one consonant) an acute can stand on the final, penultimate or antepenultimate syllable and a circumflex on the penultimate; if it is heavy, an acute can stand on the final or penultimate syllable and a circumflex on the final. A word which bears its accent as far from the word-end as permitted by the law of limitation is said to have recessive accent. Most verbal forms are accented this way. A handful of verbal categories have specific accents, for example the strong aorist participle, accented finally, e.g. εἰπὼν ‘having said’. Nouns, adjectives, adverbs and pronouns have a particular accent as part of their lexical specification. A further rule, known as the *sōtēra* rule, stipulates that if the accent falls on a long vowel in the penultimate syllable and the final vowel is short, the accent is circumflex (thus σωτήρα, not \*σωτήρα). The accent is also circumflex if it falls on a contracted final syllable. For all of the foregoing purposes the diphthongs *ai* and *oi*, when word-final, count as short rather than long vowels except in optatives and locative adverbs.

Certain types of word, including indefinite pronouns, adjectives and adverbs, and some sentence particles and forms of the verbs εἰμί ‘be’ and φημί ‘say’, are atonic enclitics which cohere so closely to a preceding word as to form a single unit for accentual purposes. An oxytone word followed by an enclitic retains its acute accent (thus ἀγαθὸς τις ἄνθρωπος vs ἀγαθὸς ἄνθρωπος). In some other cases either the preceding word or the enclitic may gain a secondary accent (e.g. ἄνθρωπός τις).

The remarks made thus far concern the accentuation of Attic and the Koine, until around the second or third century AD. Very little is known about the accentual systems of other dialects. Ancient grammarians and papyri concur that in Lesbian the accent was recessive; thus Lesb. θεῖοσι = Att. θεοῖς ‘gods (dat. pl.)’. Grammarians and papyri also suggest that in Doric the *sōtēra* rule did not apply, and that final *-ai* and *-oi* counted as long in nominative plurals. Chadwick (1992) argues that Thessalian had a word-initial stress accent. Previously, editors of epigraphic texts have accented dialect material (except Lesbian) in the manner of Attic. It is common today to omit accents in dialect texts.

Between the second and third centuries AD the accent became one of stress, as it is in Modern Greek.

## MORPHOLOGY

In both nominal and verbal morphology Greek inherited three values of the category number: singular, dual and plural. In Mycenaean whenever two items are recorded the dual is used, for example in the list of banqueting equipment from Pylos (PY Ta 641.2) *dipahe medzohē triōwehe* ‘two larger three-handled goblets’, where the noun and both adjectives are dual. In Homer and in classical Attic the dual is used (i) for “natural pairs” (such as eyes, hands, etc.); and (ii) for “accidental pairs” when the numeral δύο ‘two’ (itself a dual form) or the word ζεύγος ‘pair’ is used in conjunction; but even in these cases the plural is frequently used. Ionic had lost the category early, and its retention in Attic was considered parochial. It was therefore eliminated in Great Attic and the Koine.

### Nominal morphology

There are three main nominal declensional types: the *a*-stems (the Greek “first declension”) continue the PIE \*-eh<sub>2</sub> and \*-ih<sub>2</sub> types, the “second declension” the IE *o*-stems, and the “third declension” the various IE athematics.

TABLE 5.3 FORMS OF THE FIRST DECLENSION (*A*-STEMS) IN CLASSICAL ATTIC

		-ā type (f.)		-ǎ type (f.)		m.	
sg.	nom.	χώρᾱ-	γνώμᾱ-	ὑγίειᾱ-	δόξᾱ-	νεανιᾱ-	κριτᾱ-
	voc.	χώρᾱ	γνώμη	ὑγίειᾱ	δόξᾱ	νεανιᾱς	κριτῆς
	acc.	χώρᾱν	γνώμην	ὑγίειᾱν	δόξᾱν	νεανιᾱν	κριτῆν
	gen.	χώρᾱς	γνώμης	ὑγίειᾱς	δόξῆς	νεανίου	κριτοῦ
	dat.	χώρᾱ	γνώμῃ	ὑγίειᾱ	δόξῃ	νεανίᾱ	κριτῆ
du.	nom./acc./voc.	χώρᾱ	γνώμᾱ	ὑγίειᾱ	δόξᾱ	νεανιᾱ	κριτᾱ
	gen./dat.	χώρᾱιν	γνώμᾱιν	ὑγίειᾱιν	δόξᾱιν	νεανιᾱιν	κριτᾱῖν
pl.	nom., voc.	χώρᾱι	γνώμᾱι	ὑγίειᾱι	δόξᾱι	νεανιᾱι	κριτᾱί
	acc.	χώρᾱς	γνώμᾱς	ὑγίειᾱς	δόξᾱς	νεανιᾱς	κριτᾱς
	gen.	χώρᾱων	γνώμᾱων	ὑγίειᾱων	δόξᾱων	νεανιᾱων	κριτᾱῶν
	dat.	χώρᾱις	γνώμᾱις	ὑγίειᾱις	δόξᾱις	νεανιᾱις	κριτᾱίς

The first declension (Table 5.3) has two sub-types. The first, deriving from the \*-eh<sub>2</sub> declension, has -ā- throughout the sg., becoming -η- in Ionic and in Attic except after ε, ι or ρ. The second, deriving from the \*-ih<sub>2</sub> type, has -ǎ- in the nom., voc. and acc. sg., and -ᾱ- (becoming -η-) in the gen. and dat. sg. Masculines of the first declension have nom. sg. in -ᾱς (becoming -ης) by analogy with the second declension and gen. sg. in -ου (in Attic-Ionic), imported from the second declension. (Other dialects have forms deriving from -āo, where the -ā- is the stem and the -o has been imported from the second declension original form in -oio.) Acc. pl. -ᾱς is from -ans, which is retained in some dialects, including Mycenaean; Lesbian has -αις by the normal treatment of word-final -ns in that dialect. The gen. pl. of all *a*-stems is originally in -āhōn (preserved in Mycenaean), which contracts in Attic-Ionic to -ῶν and in West Greek to -ᾱν. There are no neuters of the first declension.

TABLE 5.4 FORMS OF THE SECOND DECLENSION (O-STEMS) IN CLASSICAL ATTIC

			m./f.	n.
sg.	nom. voc. acc. gen. dat.		λογο- 'speech'	ζυγο- 'yoke'
			λόγος	ζυγόν
			λόγε	ζυγόν
			λόγον	ζυγόν
			λόγου	ζυγοῦ
		λόγῳ	ζυγῶ	
du.	nom./acc./voc. gen./dat.		λόγῳ	ζυγῶ
			λόγῳιν	ζυγοῖν
pl.	nom., voc. acc. gen. dat.		λόγοι	ζυγά
			λόγους	ζυγά
			λόγων	ζυγῶν
			λόγοις	ζυγοῖς

The second declension (Table 5.4) consists largely of masculines and neuters. There are a handful of feminines (e.g. ἡ ὁδός 'way') which decline in the same way as the masculines. The gen. sg. is originally in *-oio*, which is preserved in Mycenaean and Homer. Loss of the intervocalic glide gave *-oo* (perhaps to be reconstructed in some Homeric forms), which contracted in Attic-Ionic to *-ou*. Acc. pl. *-ους* is from *-ons*, retained in some dialects, including Mycenaean; Lesbian has *-οις*. In Attic, stems ending in *-o* and *-e* undergo contraction (e.g. νοῦς 'mind' = νόος, ὀστοῦν 'bone' = ὀστέον). Stems in *-ē* are affected by twin processes of quantitative metathesis and prevocalic shortening, which cause both *\*-ēo-* and *\*-ēō-* to become *-εω-*. Thus Proto-Greek *\*lāwos* > *\*lēos* > λεώς. This type is known as the "Attic declension" because in the Koine it is replaced by the (non-Attic-Ionic) common Greek forms such as λαός.

TABLE 5.5 UNDERLYING ENDINGS OF THIRD DECLENSION, AND CONSONANT STEM FORMS IN CLASSICAL ATTIC

				m./f.		n.	
sg.	nom. voc. acc. gen. dat.	underlying ending	πατερ- 'father'	ἔρωτ- 'love'	ποιμεν- 'shepherd'	σωματ- 'body'	
			-s	πατήρ	ἔρωξ	ποιμήν	σῶμα
			(zero)	πάτερ	ἔρωξ	ποιμήν	σῶμα
			-η > -a	πατέρα	ἔρωτα	ποιμένα	σῶμα
			-os	πατρός	ἔρωτος	ποιμένος	σώματος
		-i	πατρί	ἔρωτι	ποιμένι	σώματι	
du.	nom./acc./voc. gen./dat.	-e	πατέρε	ἔρωτε	ποιμένε	σώματε	
		-oin	πατέροιν	ἔρώτοιν	ποιμένοιν	σώματοιν	
pl.	nom., voc. acc. gen. dat.	-es	πατέρες	ἔρωτες	ποιμένες	σώματα	
		-ης > -as	πατέρας	ἔρωτας	ποιμένας	σώματα	
		-ōn	πατέρων	ἔρώτων	ποιμένων	σώμάτων	
		-si	πατράσι	ἔρωσι	ποιμέσι	σώμασι	



The third declension consists of many sub-types. Only an overview can be given here; consult a reference grammar such as Smyth and Messing 1956 for a complete catalogue of forms. The basic underlying endings are given in Table 5.5, but because they are joined directly to the stem with no intervening vowel, if the stem ends in a consonant it undergoes various changes. Variations in stem ablaut must also be taken into account. Where the stem ends in a consonant other than *-s* or *\*-w* (also Table 5.5) the oblique cases (other than the dat. pl.) generally show the underlying form of the stem. An important archaic type is reflected in the noun *πατήρ* ‘father’, which shows the *e*-grade of the stem in the nom., voc., acc. sg., in the dual, and in the nom., voc., acc. and gen. pl., but the zero-grade in the other cases. The nom. sg. of consonant stems regularly ends in *-s*, but when the stem ends in a liquid or nasal it has no termination, and the preceding vowel, if short, lengthens (as in both *πατήρ*, stem *πατερ-*, and *ποιμήν* ‘shepherd’, stem *ποιμεν-*). In the acc. sg. the ending *\*-η* (< PIE *\*-η*) surfaces as *-α*, and in the acc. pl. *\*-ηs* as *-ᾶς*.

Neuter C-stems are endingless in the nom., voc. and acc. sg. The stem-final consonant is dropped if it is one which is not permitted at word-end (as in *σῶμα* ‘body’ from the stem *σῶματ-*).

When the stem-final consonant was *-s*, between vowels it had passed to [h] in the second millennium (hence Myc. *p<sup>h</sup>arweha* < Proto-Greek *\*p<sup>h</sup>arwesa* ‘cloths (nom./acc. pl.)’). In Attic the result undergoes contraction (Table 5.6), and it is therefore customary to treat the *s*-stems as a special type apart from the C-stems proper. There are two main sub-types. For those with a stem in *\*-es-* throughout (e.g. *Σωκράτης*), most, if not all, are masculine personal names (the one exception, *τριήρης* ‘trireme’, is probably in origin an adjective). The second major group consists of neuters with nom./voc./acc. sg. in *-os* and *\*-es-* in the remaining cases, e.g. *γένος* ‘family’.

TABLE 5.6 FORMS OF THE THIRD-DECLENSION *S*-STEMS, *I*-STEMS AND *U*-STEMS IN CLASSICAL ATTIC

		<i>s</i> -stems		<i>i</i> -stems	<i>u</i> -stems	
		m.	n.			
sg.	nom.	Σωκράτεσ- (man’s name)	γενεσ- ‘family’	πολι- ‘city’	σν- ‘pig’	πῆχρ- ‘forearm’
	voc.	Σωκράτης	γένος	πόλις	σὺς	πῆχρς
	acc.	Σωκράτη	γένος	πόλιν	σὺν	πῆχρν
	gen.	Σωκράτους	γένους	πόλεως	συός	πῆχεως
	dat.	Σωκράτει	γένει	πόλει	σὺί	πῆχει
du.	nom./acc./voc.	—	γένη	πόλει	σὺε	πῆχει
	gen./dat.	—	γενοῖν	πολέοιν	συοῖν	πῆχέοιν
pl.	nom., voc.	—	γένη	πόλεις	σὺες	πῆχεις
	acc.	—	γένη	πόλεις	σὺας, σὺς	πῆχεις
	gen.	—	γενῶν	πόλεων	συῶν	πῆχεων
	dat.	—	γένεσι	πόλεσι	συσί	πῆχεσι

The *i*-stems in Greek (Table 5.6) consist largely of feminines, though there are a few neuters (e.g. *σίναπι* ‘mustard’) which differ only in the nom. and acc. sg., which are endingless, and pl. (*σινάπι* < *σινάπεα*). The type has been extensively remodelled from

PIE. First, the forms which had the full-grade stem \*-ey- have been changed to \*-ēy- (e.g. gen. sg. \*polēyos). Loss of the intervocalic \*-y- leads to forms such as gen. sg. πόλιος (attested in Homer). In Attic we see the effects of quantitative metathesis (πόλιος > πόλειος) and prevocalic shortening (\*polēōn > πόλειον). Other dialects have generalised a stem in -i (gen. sg. πόλιος etc.).

The *u*-stems (Table 5. 6) are of two types. The more common (e.g. σῦς ‘pig’, gen. sg. σῦός) follows the PIE *-uH*-stems, having an invariant stem in *-u-* to which the C-stem endings are added. The other (e.g. πῆχυς ‘forearm’) follows the PIE *u*-stems proper with an ablauting stem in *-(e)u-*. Nouns of this type are considerably rarer, though adjectives are well represented. In nouns Attic has a gen. sg. in *-εως*, which has been imported from the *i*-stems. Other dialects (and, in adjectives, Attic) have *-εος*, which goes back to a remodelled \*-ewos.

**TABLE 5.7 UNDERLYING FORMS OF THE *-ĒU* STEMS** (forms directly attested in Mycenaean are not asterisked), **AND THE RESULTING FORMS IN CLASSICAL ATTIC**

		underlying ending	early AI	Att.
sg.	nom.	*-ēws > <i>-eus</i>	-εϋς	βασιληϋ- ‘king’ βασιλεϋς
	voc.			βασιλεϋ
	acc.	*-ēwη > *-ēwa	-ηα	βασιλέα
	gen.	<i>-ēwos</i>	-ηος	βασιλέως
	dat.	<i>-ēwi</i>	-ηι	βασιλεῖ
du.	nom./acc./voc.	*-ēwe	-ηε	βασιλεῖ
	gen./dat.	*-ēwoin	-ηοιν	βασιλέοιν
pl.	nom., voc.	*-ēwes	-ηεζ	βασιλήζ, βασιλείζ
	acc.	*-ēwηs > *-ēwas	-ηαζ	βασιλέαζ
	gen.	*-ēwōn	-ηων	βασιλέων
	dat.	*-ēwsi > <i>-eusi</i>	-εϋσι	βασιλεϋσι

Special mention must be made of the nouns in *-εϋς* (Table 5.7), mostly denominatives denoting occupations (e.g. ἵππεϋς ‘horseman’) or instruments (e.g. τομεϋς ‘knife’), but also ethnics (e.g. Μεγαρεϋς ‘Megarian’) and personal names (e.g. Ὀδυσσεϋς). The type is unique to Greek. The underlying forms have stems in *-ēw-* to which the standard C-stem endings are added. When the stem comes into contact with a consonant – i.e. in the nom. sg. and dat. pl. – Osthoff’s Law shortens the stem vowel to give nom. *-eus* and dat. *-eusi*. Other forms have intervocalic *-w-*, which is lost early in Attic-Ionic. In Attic the paradigm is further affected by quantitative metathesis (e.g. *-ēos* > *-eōs*) and prevocalic shortening (gen. pl. *ēōn* > *-eōn*).

Turning now to the inventory of case forms, the eight cases of PIE (nom., voc., acc., gen., dat., abl., instr. and loc.) are reduced in the dialects of the first millennium to five. Of these the nom., voc. and acc. directly continue their PIE counterparts. The PIE abl. was distinct only in the singular of the *o*-stems and was elsewhere identical with the gen. in the singular and the dat. in the plural. In Proto-Greek the abl. has been subsumed completely in the gen. (By a simplification of prepositional case government ablative prepositions have come to govern the dative rather than genitive in Arcado-Cypriot and probably Mycenaean (Morpurgo Davies 1966), but this does not imply an ablative-dative

syncretism in these dialects, since the genitive continues ablative functions in non-prepositional constructions. See also Thompson (1998, 2000, 2014) for further discussion of the Mycenaean evidence.)

The PIE dat., loc. and instr. collapse in a complex series of mergers which are underway at the time of the Mycenaean tablets. The first-millennium dialects have a case conventionally called “dative” which carries the functions of the PIE dat., loc. and instr. In the *a*- and *o*-stem declensions this continues, in the singular, the forms of PIE dat. sg. in *\*-āi* and *\*-ōi*. (In some dialects this undergoes later shortening, e.g. in Arc., Boe. and NWGr. *-oi*. The Mycenaean spellings *-a* and *-o* are ambiguous between the PIE dat. and loc. in *-ai*, *-oi*, but in light of the developments in the other dialects it is generally assumed that they represent *-āi*, *-ōi*.) The PIE loc. sg. *-ai*, *-oi* and *-ei* of these declensions survive only in adverbs (e.g. ἐκεῖ ‘over there’) and isolated forms such as οἶκος ‘at home’. In the third declension, however, the dialects of the first millennium continue not the PIE dat. in *\*-ey* but the loc. in *\*-i*. In Mycenaean we see a transitional stage, with forms spelt <e> = dat. *-ei* and <i> = loc. *-i* used interchangeably with both dative and locative function. The instr. sg. of all three declensional classes would be spelt in an identical way with the corresponding dat. and loc. It is thus impossible to tell whether a separate instr. sg. still existed.

In the plural no Greek dialect preserves forms of the PIE dat. in *\*-b<sup>h</sup>os* (cf. Lat. *-bus*). Already in Mycenaean the endings of the loc. pl. *-āsi* > *-āhi*, *-oisi* > *-oihi*, *-si* (remodelings of the PIE forms in *\*-āsu*, *\*-oy<sup>s</sup>u*, *\*-su*) have replaced those of the dat. Mycenaean does, however, preserve a distinct instr. pl. with endings *-a-pi* = *-āp<sup>hi</sup>i*, *-o* = *-ois* and *-pi* = *-p<sup>hi</sup>i*. The *-p<sup>hi</sup>i* endings of the *a*- and *C*-stems have disappeared from first-millennium Greek, except in Homer, where they have been reanalysed as generic oblique case markers and extended into the singular and the *o*-stems. Although these are the only plural exponents of instrumental force in Mycenaean they can also have locative and even pure dative force: *Sp<sup>h</sup>agiamp<sup>hi</sup>i* (instr. pl.) occurs in parallel with *Helehi* (loc. sg.; both are place names), and *k<sup>h</sup>itomp<sup>hi</sup>i* describes cloth ‘for khitons’. In Mycenaean, then, we seem to have a transitional stage where remnants of the morphology of dat. and loc. remain in free variation to express a syncretic dat.-loc., with which the instr. is also beginning to fall together.

In Attic-Ionic and in Lesbian the *o*-stem dat. pl. *-οισι* simply continues the loc. pl. Attic had generalised a shortened form in the definite article, τοῖς (presumably a prevocalic sandhi variant) by the beginning of the fifth century, and the same ending was standardised in nouns and adjectives by 420 BC. Other dialects have *-οις* from the earliest times, which is presumed to continue the instr. pl. In the *a*-stems early Attic and Ionic inscriptions have *-ησι* (also *-āσι* in Attic), remodelled by analogy with *-οισι* to give *-ησι/-ασι* (written *-ησι*, *-āσι*), but in Attic these were replaced by *-αισι*, and, by 420 BC, by the short form *-αις*. (Lesbian has *-αισι* from the earliest times by analogy with *-οισι*, and the other dialects have *-αις* by analogy with *-οις*.)

### Adjectives and adverbs

Adjectives follow the nominal declension. The most common type has forms of the second declension for the masculine and neuter and of the first for the feminine, e.g. σοφός, σοφή, σοφόν ‘wise’, and φίλος, φιλή, φίλιον ‘friendly’. Some, however, have no separate feminine forms (e.g. ἀθάνατος, *-ον* ‘immortal’).

*C*-stem formations are most common in participles, e.g. pres. ptc. λέγων, gen. sg. λέγοντος ‘speaking’; the n. nom./acc. sg. is endless λέγων (< \*legont), the pl. in *-a*

(λέγοντα), and the f. follows the first declension (λέγουσα < \*legontya). Similarly weak aor. pterp. λύσας, λύσασα, λῦσαν (m./n. stem λυσαντ-), and perf. λελυκώς, λελυκυῖα, λελυκός, the m./n. originally an *s*-stem (Mycenaean preserves n. pl. forms in *-oha*) but recharacterised as a *C*-stem, λελυκοτ-.

Although not a productive class, *u*-stem adjectives are well represented, following the declension of πῆχυν in the m. and n., but with gen. sg. -εος. The f. follows the first declension and is built synchronically to a stem in -ει- (e.g. ἡδύς, ἡδεῖα, ἡδύ ‘sweet’, m./n. gen. sg. ἡδέος). *s*-stem adjectives (e.g. εὐγενής ‘well-born’) have stems in *-es-* with loss of the *-s-* between vowels, as in nouns like Σωκράτης. They have no separate f. The n. nom./acc. sg. is endingless (εὐγενές), and pl. in \**-esa* > -εα > -η (εὐγενῆ).

The adjective also has comparative and superlative degrees. There are two different methods of formation. In the first the suffixes *-yon-* or *-ion-* (comparative) and *-ιστος* (superlative) are added directly to the root (not the stem) of the adjective, originally in the *e*-grade: μέγας ‘big’ gives comparative μείζων (-ει- by an internal development in Attic for μέζων of other dialects; -ζ- < \**-gy-*) and superlative μέγιστος. ἡδύς ‘sweet’ gives comparative ἡδίον and superlative ἡδιστος. αἰσχρός (root αἰσχ-) gives comparative αἰσχίον and superlative αἰσχιστος. The original comparative suffix was \**-yos-* (cf. Lat. *-ior*), an *s*-stem attested in Myc. *medzohes* ‘bigger (m./f. nom. pl.)’, and preserved in the Attic alternative m./f. acc. sg. μείζω < \**-oha*, nom. and acc. pl. μείζους < \**-ohes*, \**-ohas*, and n. nom./acc. pl. μείζω < \**-oha*. In the zero-grade it was compounded with an indefinite nominalising suffix *-on-* to give *n*-stem *-iov-*. Forms such as μείζων are a conflation of the two. The superlative *-ιστος*, a regular first/second declension adjective, is a compound of the same zero-grade \**-is-* and a definite nominalising suffix \**-to-*.

The second formation adds *-τερος* (comparative) and *-τατος* (superlative) to the (m./n.) stem. Both are first/second declension adjectives. This *-τερος* was originally contrastive in force rather than comparative, a meaning preserved in formations such as δεξιτερος ‘right-hand’ (contrasting with σκαιός ‘left-hand’) and ἀριστερος ‘left-hand’ (contrasting with δεξιός ‘right-hand’). In Mycenaean, *wa-na-ka-te-ro* = *wanakteros* means ‘of the royal (rather than of another) type’. The same suffix is also seen in the possessives ἡμέτερος ‘our’ and ὑμέτερος ‘your’. The superlative *-τατος* appears to be a contamination of the PIE absolutive suffix \**-(t)ṛH-* (Lat. *ultimus*) by *-ιστο-*. Examples of paradigmatic comparatives and superlatives are δεινότερος, δεινότατος from δεινός ‘strange’, γλυκύτερος, γλυκύτατος from γλυκύς ‘sweet’, μελάντερος, μελάντατος from μέλας ‘black’ (stem μελαν-). Stems in *-o-* lengthen the theme vowel if the preceding syllable is light, e.g. σοφότερος, σοφότατος from σοφός ‘wise’ (instead of \*σοφότερος, \*σοφότατος). Metanalysis of formations such as ἀληθέσ-τερος ‘more true’ gives rise to *-έστερος* and *-έστατος* (so εὐδαιμονέστερος built to εὐδαιμών ‘fortunate’), and of ἀχαρίσ-τερος ‘more ungracious’ (< \**ak<sup>h</sup>arid-tero-*) to *-ίστερος*, *-ίστατος* (so λαλίστερος built to λάλος ‘talkative’). Similarly, forms built regularly to adverbs such as παλαιότερος ‘older’ (to πάλαι ‘long ago’) instead of the positive grade of the adjective (παλαιός ‘ancient’) cause some adjectives in *-αιος* to form comparatives and superlatives in *-αίτερος* and *-αίτατος* rather than expected *-αιότερος*, *-αιότατος* (e.g. ἡσυχᾶος, ἡσυχαιότερος, ἡσυχαιάτατος ‘quiet(er, -est)’), and conversely some forms in *-ος* to give *-αίτερος*, *-αίτατος* rather than expected *-ότερος*, *-ότατος* (e.g. μέσος ‘middle’ gives μεσαίτερος, μεσαίτατος).

Besides these regular formations there are many irregularities caused by stem ablaut and other phonetic changes (πολύς ‘much’ gives πλείων/πλέων, πλείστος) and by suppletion (ἀγαθός ‘good’ has both βελτίων, βέλτιστος and ἀμείνων, ἄριστος).

Adverbs are regularly derived from adjectives by the addition of the suffix *-ως* to the (m./n.) stem, thus, e.g., φίλως ‘friendly’ from φίλος, ταχέως ‘swiftly’ from ταχύς;

but many are fossilised case forms (e.g. ἐκεῖ ‘over there’ with *o*-stem loc. sg. ending). The comparative and superlative of the adverb are identical with the n. nom./acc. sg. comparative and pl. superlative of the corresponding adjective (e.g. σοφῶς ‘wisely’ has σοφώτερον, σοφώτατα).

### Pronouns

The definite article is in origin a pronoun built to the PIE demonstrative stems \*so (m. nom. sg. is endingless, *ó*; f. *ἦ* < *ā*, preserved outside Attic-Ionic) and \*to (other forms; West Greek preserves m. and f. nom. pl. in τοι, ται, while the other dialects have οἱ, αἱ by analogy with the sg.). In Attic it is obligatory with nouns of definite reference, including personal names and abstracts, but its use as an article is evidently a late development. It is entirely absent from Mycenaean, and although in Homer it has uses which approach those of later Greek, it still functions as an independent pronoun with demonstrative, anaphoric, contrastive or relative force. The use as a relative (compare English *that*) survives in the dialects and in Attic verse.

As a demonstrative the same pronoun survives in the composite ὄδε, ἦδε, τόδε, which has a strongly deictic force ‘this one (over here)’. Other dialects have similar forms with different deictic particles, e.g. Arc. (h)οῦν. It contrasts with ἐκεῖνος, ἐκείνη, ἐκεῖνο ‘that one (over there)’. The pronoun οὗτος, αὗτη, τοῦτο ‘this’ has more anaphoric than deictic force.

TABLE 5.8 FORMS OF THE PERSONAL PRONOUNS (enclitic forms are bracketed)

		First person	Second person	Third person/reflexive
sg.	nom.	ἐγώ	σύ (also voc.)	—
	acc.	ἐμέ (με)	σέ (σε)	ἔ (έ)
	gen.	ἐμοῦ (μου)	σοῦ (σου)	οὔ (οῦ)
	dat.	ἐμοί (μοι)	σοῖ (σου)	οῖ (οί)
du.	nom./acc.	νό	σφώ	(σφωε)
	gen./dat.	νῶν	σφῶν	(σφωῖν)
pl.	nom.	ἡμεῖς	ὑμεῖς (also voc.)	σφεῖς
	acc.	ἡμᾶς	ὑμᾶς	σφέας, σφᾶς (σφεας)
	gen.	ἡμῶν	ὑμῶν	σφέων, σφῶν (σφεων)
	dat.	ἡμῖν	ὑμῖν	σφίσι (σφισι)

The forms of the first and second person pronouns are shown in Table 5.8. The singular has both emphatic accented and unemphatic enclitic forms. The second person nom. sg. should be \*tu (cf. Lat. *tū*), and this is the form preserved in WGr. τῦ. The oblique forms are \*twe etc., which give Gr. σέ etc.; and the other dialects have levelled the σ- to the nom. There is in Attic no third person pronoun *per se*. For the oblique cases the pronoun αὐτός, αὐτή, αὐτό is pressed into service, but this is properly an intensive pronoun (corresponding to Lat. *ipse*), and its nom. cannot be used as a third person pronoun. In the combination ὁ αὐτός it means ‘the same’. In Homer the pronoun ἔ is sometimes reflexive, more often anaphoric (when it is always enclitic). In the reflexive use it is often strengthened with the appropriate case of αὐτός. In Attic ἐ- has become an indeclinable prefix conjoined to αὐτός to form the standard reflexive pronoun ἐαυτόν, ἐαυτήν, ἐαυτό

‘himself’; so also ἐμ- and σε- in ἐμαυτόν and σεαυτόν ‘myself’ and ‘yourself’. The only forms of εἶ which are regularly used in Attic prose are the dat. sg. οἷ and pl. σοφίσι, which are used as indirect reflexives (to refer in a subordinate clause to the subject of the matrix clause). Mycenaean uses the dat. *sp<sup>h</sup>ehi* (= σοφίσι) as an anaphoric pronoun in the phrase *meta-k<sup>w</sup>e sp<sup>h</sup>ehi* ‘and with them’. The enclitic acc. sg. μιν ‘him/her’, seen in Homer, is also attested in Mycenaean in the phrase *dāmos de min p<sup>h</sup>āsi . . . hek<sup>h</sup>ehen* ‘but the dāmos says that she has . . .’.

Of the PIE pronominal stems \*k<sup>w</sup>i- and \*yo- both survive into Greek. The latter is the basis of the relative pronoun ὅς, ἣ, ὅ. The former is remodelled as an *n*-stem τίς, *n*. nom. sg. τί, gen. sg. τίνοϛ, and becomes the standard interrogative pronoun ‘who?’ In an enclitic form it serves as the indefinite pronoun. Compounded with the relative pronoun ὅς, it gives the indefinite relative and indirect interrogative ὅστις, ἣτις, ὅτι.

### Verbal morphology

The verb expresses the categories of person (first, second, third), number, time reference (past, present, future), aspect (durative, punctual, stative), mood (indicative, subjunctive, optative, imperative) and voice (active, middle, passive). Non-finite forms include infinitives and participles. In the indicative, time reference and aspect combine to give the seven tenses of traditional grammars, present, future, imperfect (past durative), aorist (past punctual), perfect (originally present stative), pluperfect (originally past stative) and future perfect.

The difference between durative and punctual aspect is encoded in different stems, traditionally labelled “present” and “aorist”. In some verbs the present stem seems to be basic, and the aorist stem is derived, usually by the addition of *-s-* (e.g. λύ-ω ‘loose’, aor. ἔ-λυ-σ-α). In others the aorist stem seems basic, and the present stem is derived by one of several means including ablaut (aor. ἔ-λιπ-ον, pres. λείπ-ω ‘leave’); suffixation with *-σκε/ο-*, originally an iterative/durative suffix (e.g. aor. ἔ-παθ-ον, pres. πάσχω < \*πάθ-σχω ‘suffer’); reduplication (e.g. aor. ἔ-στη-ν, pres. ἵ-στη-μι < \*si-stā-mi); or addition of a nasal suffix (e.g. aor. ἔ-τεμ-ον, pres. τέμ-ν-ω ‘cut’; aor. ἤμαρτ-ον, pres. ἀμαρτ-άν-ω ‘err’) and/or infix (e.g. aor. ἔ-μαθ-ον, pres. μα-ν-θ-άν-ω ‘learn’). The difference between the two types is sometimes explained as motivated by the inherent atelicity (underived present stem) or telicity (underived aorist stem) of the verbal action, but the correlation is not perfect. In yet other verbs both stems are derived (e.g. pres. αὐξ-άν-ω ‘increase’, aor. ἠῦξ-ησ-α). Especially common is the large class where the present stem is derived (often from a nominal stem) by the suffix *\*-ye/o-*, the aorist stem with the *\*-s-* suffix (e.g. pres. φυλάττω/φυλάσσω ‘guard’ with *-ττ-/-σσ-* < *\*-ky-*, aor. ἐφύλαξα = ἐ-φύλακ-σ-α; pres. κόπτω ‘strike’ with *-πτ-* < *\*-py-*, aor. ἔκοψα = ἔ-κοπ-σ-α; pres. ἀγγέλλω ‘announce’ with *-λλ-* < *\*-ly-*, aor. ἠγγεῖλα < \*āngel-s-a). There are many verbs where the different stems are provided by suppletion (e.g. pres. φέρω ‘bear’, aor. ἦνεγκα; pres. ὁράω ‘see’, aor. εἶδον), and others where the operation of sound change on a regularly derived paradigm creates the appearance of suppletion (e.g. aor. ἔ-μολ-ον, pres. βλώσκω < \*m<sup>l</sup>h<sub>3</sub>-ské/o-).

The distinction between present and past time reference is encoded partially in the endings (which also encode for person, number and voice) and partially in the presence or absence of the augment. In the dialects of the first millennium the augment ἐ- is a compulsory marker of past time reference prefixed to the stem of indicatives. (In roots starting with a vowel it surfaces as lengthening of the initial vowel, and is called the temporal, as opposed to syllabic, augment.) In Homer, as in the R̥g Veda, the augment is optional, being present or absent at the convenience of the poet. Absence of the augment was

formerly held to be a poetic feature. Its almost total absence from Mycenaean, however – there are one or two possible examples – can scarcely be so explained.

**TABLE 5.9 THEMATIC AND ATHEMATIC PRIMARY AND SECONDARY ENDINGS OF THE ACTIVE VOICE**

		Thematic		Athematic			
		λύω ‘loose’		τίθημι ‘place’		εἰμί ‘be’ (irregular)	
		Primary (present)	Historic (imperfect)	Primary (present)	Historic (imperfect)	Primary (present)	Historic (imperfect)
sg.	1	λύ-ω	ἔ-λυ-ον	τί-θη-μι	ἔ-τί-θη-ν	εἰμί	ἦ, ἦν
	2	λύ-εις	ἔ-λυ-ες	τί-θη-ς	ἔ-τί-θη-ς	εἶ	ἦσθα
	3	λύ-ει	ἔ-λυ-ε	τί-θη-σι	ἔ-τί-θη	ἔσ-τί	ἦν
du.	2	λύ-ετον	ἔ-λύ-ετον	τί-θε-τον	ἔ-τί-θε-τον	ἔσ-τόν	ἦστον
	3	λύ-ετον	ἔ-λυ-έτην	τί-θε-τον	ἔ-τι-θέ-την	ἔσ-τόν	ἦστην
pl.	1	λύ-ομεν	ἔ-λύ-ομεν	τί-θε-μεν	ἔ-τί-θε-μεν	ἔσ-μέν	ἦμεν
	2	λύ-ετε	ἔ-λύ-ετε	τί-θε-τε	ἔ-τί-θε-τε	ἔσ-τέ	ἦτε
	3	λύ-ουσι	ἔ-λυ-ον	τι-θέ-ᾱσι	ἔ-τί-θε-σαν	εἰσί	ἦσαν

In the indicative the present stem combines with the so-called primary endings to form the present tense and with the secondary or historic endings (and augment) to form the imperfect tense (past durative) – see Table 5.9. Two different formations must be distinguished, one thematic (presents in -ω), one athematic (presents in -μι). The latter are no longer productive in Greek, and tended to be replaced by thematic formations.

From the derivational point of view the historic endings are in fact basic, and several of the primary forms are transparently derived from them by the addition of the \*-i “hic et nunc” deictic marker. The 1 sg. historic -ν goes back to \*-m, from which primary -μι is derived. In thematic -ον the -ο- is the thematic vowel. The thematic ending -ω is from unrelated \*-oh<sub>2</sub>, in which the -ο- is again the thematic vowel.

The 2 sg. historic -(ε)ς is original (-ε- is the thematic vowel). The expected primary endings would be in \*-si, where the -s- would be lost (via [h]) between vowels. The resulting forms \*luei and \*tit<sup>h</sup>ēi had the -s- restored by analogy with the secondary endings: in the thematic class it is added to the end to give λύεις, while the athematic class borrows the entire secondary ending.

In the 3 sg. the athematic secondary ending was \*-t, which was lost at the word-end. Athematic -σι derives from \*-ti by the normal assibilation \*-ti > -si in East Greek dialects. The original -τι is preserved in ἐστί and in West Greek (τιθητι etc.). Thematic -ει cannot go back to \*-eti, which would remain in West Greek and give \*-εσι in East Greek. The standard explanation of -ει is that it results from the four-part analogy ἔλυες : ἐλύεις :: ἔλυε : X, X = ἐλύει. Sihler (1995: 462) instead endorses the view that the thematic historic 3 sg. was endless \*-e, and that -ει derives straightforwardly by the addition of *i*-deictic.

Athematic root presents like τίθημι show stem ablaut with the full grade (θη- < \*d<sup>h</sup>eh<sub>1</sub>-) in the sg. and the zero-grade (θε- < \*d<sup>h</sup>h<sub>1</sub>-) in the pl. In the verb ‘be’ the sg. *es-* goes back to \*h<sub>1</sub>es-, the pl. *es-* presumably to \*h<sub>1</sub>s- (cf. Ved. *sánti*). The athematic secondary 3 pl. ending was originally \*-nt. The imperfect of ‘be’ \*e-h<sub>1</sub>s-ent > \*ēhen, whence WGr. ἦν, a form which was replaced in Attic-Ionic by ἦσαν with an ending from the weak aorist.

Original ἦν was redeployed in Attic as a 3 sg. Outside Attic-Ionic the other athematic verbs have historic 3 pl.  $-v < *nt$  (imperfect ἔτιθεν etc.). Attic-Ionic has made the same replacement as in ἦσαν. Thematic primary  $-ουσι$  shows the expected East Greek development from  $*-o-nti$  (with  $-o-$  as the thematic vowel); West Greek preserves  $-οντι$ . Athematic  $-ᾶσι$  is peculiar to Attic. West Greek has  $τιθεντι, διδοντι, ισταντι$  etc.

In verbs with stems ending in  $-ε-$ ,  $-α-$  and  $-ο-$  in Attic and in some other dialects the stem vowel and ending coalesce in a process known as contraction, e.g.  $ποιέω > ποιῶ, ὀράομεν > ὀρῶμεν, δηλόει > δηλοῖ$ . In other dialects the vowels remain uncontracted. In yet others – particularly Aeolic – these conjugate as athematics.

TABLE 5.10 ENDINGS OF THE STRONG AND WEAK AORIST

		Strong aorist πειθω 'persuade'	Weak aorist λύω 'loose'
sg.	1	ἔ-πιθ-ον	ἔ-λυ-σ-α
	2	ἔ-πιθ-εσ	ἔ-λυ-σ-αs
	3	ἔ-πιθ-ε	ἔ-λυ-σ-ε
du.	2	ἐ-πίθ-ετον	ἐ-λύ-σ-ατον
	3	ἐ-πίθ-έτην	ἐ-λυ-σ-άτην
pl.	1	ἐ-πίθ-ομεν	ἐ-λύ-σ-αμεν
	2	ἐ-πίθ-ετε	ἐ-λύ-σ-ατε
	3	ἔ-πιθ-ον	ἔ-λυ-σ-αν

The aorist stem forms only one indicative, the *aorist tense* (Table 5.10), a punctual preterite which corresponds in meaning not only to English 'I did' but also 'I have done' and 'I had done'. There are two main formations, known as the strong (or second) and weak (or first) aorists. Most verbs follow one or the other formation, but some have both in different senses (e.g. ἵστημι 'I place' has a transitive weak aorist ἔστησα 'I placed' but intransitive strong aorist ἔστην 'I stood'), and others follow the weak formation in the sg. and the strong in the du. and pl. (e.g. τίθημι has ἔθηκα but ἔθεμεν; the 3 pl. is either weak ἔθηκαν or strong ἔθεσαν). The strong aorist uses the thematic secondary endings; it therefore differs from the imperfect only in the form of the stem. The weak aorist continues the PIE  $*-s-$  aorist, which was originally athematic but has been remodelled in Greek so that it effectively has a theme vowel  $-a$ . Additionally, there are root aorists of the type ἔβλην 'threw'  $< *e-g^wleḥ_1-$ ; these take the athematic secondary endings.

PIE had no *future tense*. The Greek future derives from the PIE desiderative suffix  $*-h_1s-$ , which was added directly to the root and originally formed athematics, but has come to take the thematic primary endings. After stems ending in a stop the laryngeal seems to have been lost, giving rise to forms such as λείψω 'I will leave' = λείπ-σ-ω. This is the regular formation in Attic. After stems ending in a resonant, however, the laryngeal was not lost, giving a suffix  $*-es-$  whose  $-s-$  was invariably intervocalic and therefore lost (via [h]). This gives rise to the so-called *liquid* or *contracted futures* such as βαλῶ = βαλέω from βάλλω (root βαλ-). The so-called *Attic futures* are an analogical extension of the liquid type to stems in which they do not properly belong and which in most dialects form sigmatic futures. Despite the name they are not, in fact, restricted to Attic. The so-called *Doric future* (e.g. κλεψεω from κλέπτω 'steal') is a blend of both types which has become standard in West Greek.



In addition to the active, Greek inherited from PIE a *middle voice*, originally denoting an action in which the agent was particularly involved, either reflexively (λούω ‘wash’ is transitive; λούομαι is ‘I wash myself’) or with specialised meaning: so, for example, while λύω means ‘I loose’, λύομαι means ‘I ransom’; and while φέρω means ‘I carry’, φέρομαι means ‘I win (i.e. carry off) a prize’. Many verbs have middle rather than active forms of the future, perhaps reflecting the original desiderative function of the future suffix; so, for example, ἔσομαι from the verb ‘be’. Some verbs, said to be deponent, have only middle forms, often for reasons which are no longer semantically transparent (so, for example, ἔπομαι ‘follow’). The middle came, however, to be the exponent of passive voice. The forms of the middle are given in Table 5.11; the future middle uses the thematic primary endings on the future stem.

TABLE 5.11 ENDINGS OF THE MIDDLE VOICE

		Present stem				Aorist	
		Thematic		Athematic		Strong	Weak
		λύω ‘loose’		τίθημι ‘place’			
		Primary (present)	Historic (imperfect)	Primary (present)	Historic (imperfect)		
sg.	1	λύ-ομαι	ἐ-λυ-όμην	τί-θε-μαι	ἐ-τί-θέ-μην	ἐ-πιθ-όμην	ἐ-λυ-σ-άμην
	2	λύ-ῃ	ἐ-λύ-ου	τί-θε-σαι	ἐ-τί-θε-σο	ἐ-πιθ-ου	ἐ-λύ-σ-ω
	3	λύ-εται	ἐ-λύ-ετο	τί-θε-ται	ἐ-τί-θε-το	ἐ-πιθ-ετο	ἐ-λύ-σ-ατο
du.	2	λύ-εσθον	ἐ-λύ-εσθον	τί-θε-σθον	ἐ-τί-θε-σθον	ἐ-πιθ-εσθον	ἐ-λύ-σ-ασθον
	3	λύ-εσθον	ἐ-λυ-έσθην	τί-θε-σθον	ἐ-τι-θέ-σθην	ἐ-πιθ-έσθην	ἐ-λυ-σ-άσθην
pl.	1	λυ-όμεθα	ἐ-λυ-όμεθα	τι-θέ-μεθα	ἐ-τι-θέ-μεθα	ἐ-πιθ-όμεθα	ἐ-λυ-σ-άμεθα
	2	λύ-εσθε	ἐ-λύ-εσθε	τί-θε-σθε	ἐ-τί-θε-σθε	ἐ-πιθ-εσθε	ἐ-λύ-σ-ασθε
	3	λύ-ονται	ἐ-λύ-οντο	τί-θε-νται	ἐ-τί-θε-ντο	ἐ-πιθ-οντο	ἐ-λύ-σ-αντο

The 3 sg. historic form *-to* is original. The derived primary ending ought to be *-toi*, which is in fact preserved in Arcado-Cypriot and in Mycenaean. *-ται* has been influenced by *-μαι* and *-σαι*. The same holds for the 3 pl. *-ντο*, *-νται*.

In early Greek the aorist middle is used with passive sense, but two distinctively passive aorist formations arose within the history of the language, dubbed strong and weak. The strong aorist passive conjugates like the (active) root aorists such as ἔβλην, with *-η-* throughout and athematic secondary endings. Indeed, it probably originates in aorists of this type, many of which are intransitive. In Homer, the vast majority of strong aorist passives are in fact intransitives, e.g. ἐχάρη ‘I rejoiced’. The weak aorist passive is formed with a suffix *-θη-* (of unknown origin) to which the athematic historic endings are added. It thus conjugates in an identical fashion to the strong aorist passive – the difference is simply in the stem. Originally *θη-* was added to the zero-grade of the stem (e.g. τείνω forms ἐτάθη), later to the weak aorist stem without – or even with – the *-σ-* marker. In Homer the *-θη-* aorist passive is frequently used interchangeably with the corresponding middle.

From the new aorist passives arose also two corresponding future passives. The stem is in  $-\eta-$  (“strong”) or  $-\theta\eta-$  (“weak”) followed by the future marker  $-\sigma-$  with primary middle endings. The  $-\theta\eta\sigma\mu\alpha\iota$  type is completely absent from Homer, the  $-\eta\sigma\mu\alpha\iota$  type attested in two forms only,  $\mu\eta\gamma\eta\sigma\epsilon\sigma\theta\alpha\iota$  (*Iliad* 10.365) and  $\delta\alpha\eta\sigma\epsilon\alpha\iota$  (*Odyssey* 3.187 and 19.325), suggesting that both are rather late formations. Supporting this view is the use in Mycenaean of a middle future participle with passive sense in the phrase *aleip<sup>h</sup>atei dzes(s)omenōi* ‘for an unguent to be boiled’.

The *subjunctive mood* is formed in thematic stems by lengthening the vowel of the ending. Thus subjunctive  $\lambda\acute{\upsilon}\eta\iota$ ,  $\lambda\acute{\upsilon}\omega\mu\epsilon\nu$  vs indicative  $\lambda\acute{\upsilon}\epsilon\iota$ ,  $\lambda\acute{\upsilon}\omega\mu\epsilon\nu$ . In athematic formations the vowel *e/o* was originally inserted between the stem and the ending; many of these “short-vowel” subjunctives survive in Homer and some dialects (e.g. subjunctive  $\acute{\iota}\omega\mu\epsilon\nu$  vs indicative  $\acute{\iota}\mu\epsilon\nu$ ), but the lengthened thematic forms quickly replaced them. On the use of the subjunctive, see under “Syntax” below.

The *optative mood* was formed in PIE in athematic stems by the addition of the suffix  $*-yeh_1-/*-ih_1-$  to the zero-grade of the root, the suffix appearing in the *e*-grade in the sg. and the zero-grade in the du. and pl. This is exactly what happens in Greek: from the verb ‘be’ 1 sg.  $\epsilon\acute{\iota}\eta\nu < *h_1s-yeh_1-m$ , 1 pl.  $\epsilon\acute{\iota}\mu\epsilon\nu < *h_1s-ih_1-men$ . Similarly, in verbs such as  $\tau\acute{\iota}\theta\eta\mu\iota$ , 1 sg.  $\tau\acute{\iota}\theta\epsilon\acute{\iota}\eta\nu < *d^hi-d^h_1-yeh_1-m$ , 1 pl.  $\tau\acute{\iota}\theta\epsilon\acute{\iota}\mu\epsilon\nu < *d^hi-d^h_1-ih_1-men$ , although one would expect the intervocalic  $*-y-$  to have been lost in the former, and, perhaps, the outcome of the latter to have been  $*\tau\acute{\iota}\theta\acute{\iota}\mu\epsilon\nu$ ; evidently some analogical rebuilding has taken place. Note that the endings are the athematic secondary ones (with original 3 pl.  $-ev < *-ent$  preserved even in Attic, rather than the replacement  $-\sigma\alpha\nu$ ).

In thematic formations Greek uses a suffix  $-ot-$ , which is perhaps the thematic vowel followed by the zero-grade of  $*-yeh_1-$  (the laryngeal having disappeared, except in Slavic). The endings are once again the athematic secondary ones, except in the 1 sg., which has been replaced by  $-οι\mu\iota$ , containing the athematic primary ending.

It is important to note that outside the indicative and participles (and some uses of the infinitive) the difference between so-called present and aorist formations is not one of time reference but of aspect. The “present subjunctive” is not a present-time subjunctive, nor does the “aorist subjunctive” refer to past time. Both indeed can refer to future events. “Aorist” forms, built to the aorist or punctual stem, refer to simple events, while “present” forms, built to the present or durative stem, refer to ongoing or repeated events.

As well as the *eventive* formations discussed so far, PIE possessed a *stative*, which is the origin of the Greek formation known as the *perfect tense*. Originally it denotes a state in which the subject finds itself – characteristic examples are  $\omicron\acute{\iota}\delta\alpha$  ‘I know’,  $\tau\acute{\epsilon}\theta\eta\eta\kappa\epsilon$  ‘he is dead’ (vs eventive  $\theta\eta\eta\sigma\kappa\epsilon\iota$  ‘he dies’ and  $(\acute{\alpha}\pi)\acute{\epsilon}\theta\alpha\nu\epsilon$  ‘he died’),  $\acute{\epsilon}\sigma\tau\eta\kappa\alpha$  ‘I am standing’ (vs  $\acute{\iota}\sigma\tau\alpha\mu\alpha\iota$  ‘I (go and) stand (somewhere)’). As such the stative originally stood outside the tense and mood system, and in Mycenaean we still find forms such as *t<sup>h</sup>et<sup>h</sup>uk<sup>h</sup>woha*, perf. ptcp. of  $\tau\acute{\epsilon}\upsilon\chi\omega$  ‘make’, with what from a later Greek perspective is active morphology but passive sense, ‘finished off, completed’; in fact, the meaning is ‘in a state of completion’. Similarly, in Homer  $\delta\acute{\epsilon}\delta\eta\epsilon$  means ‘is ablaze’, from  $\delta\alpha\acute{\iota}\omega$  ‘kindle’. Even by the time of Mycenaean, however, middle-passive forms of the perfect have been created, e.g. *dedemena* ‘bound’, perf. of  $\delta\acute{\epsilon}\omega$ , indicating that the perfect has begun to align itself with the voice system of the eventives. Semantically a present state can often be viewed as the result of a completed action in the past – ‘he died and therefore is dead’ – and so the perfect came to denote past action with continued relevance in the present time, ‘he has died’. In this sense it is properly speaking an eventive present perfect and has become aligned to the tense system. It is therefore no surprise that corresponding past and future forms were created, the pluperfect  $\acute{\epsilon}\tau\epsilon\theta\eta\eta\kappa\epsilon\iota$  ‘he had died’ and future perfect

λελύσομαι ‘I shall have ransomed/I shall have been released’ (these forms are almost all middle-passive, τεθνήξει ‘he will be dead/have died’ being a rare exception). In the Koine the perfect became interchangeable with the aorist as a simple past tense ‘I did’ (compare the development of the perfect in Latin, and the renewed periphrastic perfect in French).

TABLE 5.12 FORMS OF THE PERFECT AND PLUPERFECT IN CLASSICAL ATTIC

οἶδα ‘know’			λύω ‘loose’				
	Perfect	Pluperfect	Perfect		Pluperfect		
			Active	Middle	Active	Middle	
sg.	1	οἶδα	ἤδη	λέ-λυ-κ-α	λέ-λυ-μαι	έ-λε-λύ-κ-η	έ-λε-λύ-μην
	2	οἶσθα	ἤδησθα	λέ-λυ-κ-ας	λέ-λυ-σαι	έ-λε-λύ-κ-ης	έ-λε-λύ-σο
	3	οἶδε	ἤδει	λέ-λυ-κ-ε	λέ-λυ-ται	έ-λε-λύ-κ-ει	έ-λε-λύ-το
du.	2	ἴστον	ἤστον	λε-λύ-κ-ατον	λέ-λυ-σθον	έ-λε-λύ-κ-ειτον	έ-λε-λύ-σθον
	3	ἴστον	ἤστιν	λε-λύ-κ-ατον	λέ-λυ-σθον	έ-λε-λυ-κ-είτην	έ-λε-λύ-σθην
pl.	1	ἴσμεν	ἤσμεν	λε-λύ-κ-αμεν	λε-λύ-μεθα	έ-λε-λύ-κ-ειμεν	έ-λε-λύ-μεθα
	2	ἴστε	ἤστε	λε-λύ-κ-ατε	λέ-λυ-σθε	έ-λε-λύ-κ-ειτε	έ-λε-λύ-σθε
	3	ἴσασι	ἤδεσαν, ἤσαν	λε-λύ-κ-ᾶσι	λέ-λυ-νται	έ-λε-λύ-κ-ε(ι)σαν	έ-λε-λύ-ντο

Although it lacks the reduplication characteristic of the PIE stative, οἶδα ‘I know’ most closely represents the original with *o*-grade of the root in the sg., zero-grade in the du. and pl. (Table 5.12). In the forms which are productive in Greek, however, the endings have been heavily remodelled. Reduplication remained the hallmark of the perfect system. The so-called strong perfect usually has the root in the *o*-grade and no special suffix (e.g. λέλοιπα, perf. of λείπω ‘leave’). The “weak” or kappatic perfect adds -κ- to the present stem (e.g. λέλυκα, perf. of λύω); the origin of this formation is obscure, but it is already represented in Homer in verbs with stems in long vowels. The middle is an internal Greek development formed from the reduplicated stem (usually zero-grade) with primary athematic middle endings.

The pluperfect, again, is a Greek innovation, but it may have originated at a time when the perfect had a purely stative function to express a state in past time. The stem is the perfect stem with augment. The endings, apparently, are the perfect endings prefixed with -ε-, but the details are obscure.

The perfect stem, like the present and aorist stems, forms subjunctives and optatives, albeit rarely. They may be internal Greek developments (certainly perfect-stem optatives are vanishingly rare in Vedic). The active forms are built on the same model as the thematic subjunctives and optatives (λελύκω, λελύκοιμι); the passives are periphrastic (λελυμένος ὦ, λελυμένος εἶη).

Greek has a fully developed system of infinitives and participles, with present, future, aorist and perfect stems giving both active and middle forms. Additionally, the future and aorist have distinctively passive forms. In the accusative and infinitive indirect statement construction present and aorist infinitives have present and past time reference respectively, while in other constructions the difference is one of aspect only. Participles encode time relative to the action of the main verb (aorist for anterior, present for contemporary, future for posterior action).

## SYNTAX

### Word order

Having a richly articulated inflexional morphology Greek has relatively free word order. Nonetheless, in classical Attic prose we can detect a preference for SOV order (Dover 1960: 25), although this is often disrupted for pragmatic reasons (Dik 1995).

In Mycenaean we have in main clauses examples of both SOV (*krit<sup>h</sup>ēwiai onāton hek<sup>h</sup>onsi* ‘the barley-women have an *onāton* (a type of lease)’) and SVO (*Kōkalos ape-dōke elaiwon tosson Ehumēdehi* ‘Kōkalos gave so much olive oil to Eumēdēs’); and in subordinate clauses of SVO (*hote wanaks t<sup>h</sup>ēke Augēwān dāmo-ko-ro* ‘when the king appointed Augēwās as *dāmo-ko-ro*’).

Various words including sentence-connective particles and atonic, clitic pronouns originally occupied the “Wackernagel’s Law” position immediately after the first constituent of the sentence. The interplay between the requirements of SV order and Wackernagel’s Law frequently meant that clitic pronouns were often either widely separated from their governing verb or placed following the verb but out of second position. In the Hellenistic period the tension was resolved by putting the verb in first position so that clitic pronouns naturally followed in second position, and this led to a development of VS word order even in those sentences which lacked clitics (Horrocks 1990; 2010: 108–109).

### Mood, complementation and subordination

The subjunctive seems originally to have been used to refer to future events, while the optative expressed the category *irrealis*. Traces of these functions are preserved in the use of the subjunctive in the protasis of future open conditions (ἐὼν ταῦτα ποιῆς, καλῶς ποιήσεις ‘if you do this you will do well’) and of the optative in the protasis and apodosis of future remote conditions (εἰ ταῦτα ποιοίης, καλῶς ἂν ποιοίης ‘if you were to do this, you would do well’) and in wishes for the future (εἴθε ἔλθοι ‘would that he would come’). Over time, however, the subjunctive encroached on the modal uses of the optative, which became increasingly marginalised. In the Koine it survives only in fossilised expressions such as μὴ γένοιτο ‘God forbid!’.

Greek has a range of methods for embedding clauses as the complements of verbs, both finite and non-finite. Finite declarative and interrogative complements (i.e. expressions of indirect statement and question) are introduced by an overt complementiser (ὅτι, ὥς both ‘that’, εἰ ‘whether’ etc.) or an interrogative, and ordinarily have their verb in the indicative. Unlike English, Greek has no rule of sequence of tenses: the verb in the complement clause takes the tense of the verb in the equivalent direct statement or question: so διήλθε λόγος ὅτι διώκει αὐτοὺς Κύρος (Xenophon *Anabasis* 1.4.7) ‘a report spread that Cyrus was pursuing them’, lit. ‘is pursuing’ because the report was ‘Cyrus *is pursuing* them’; but after a past tense matrix verb (traditionally called *historic sequence*) the verb in the subordinate clause may become optative, retaining, so far as possible, the tense of the direct form (in violation of the normal principle that the optative does not express time reference but only aspect): so ἔγνωσαν ὅτι κενὸς ὁ φόβος εἶη (Xenophon *Anabasis* 2.2.21) ‘they realised that their fear was (lit. is) groundless’ with εἶη for ἐστί. Thus, Greek is often said to have a rule of *sequence of mood* rather than of tense. This sequence-of-mood rule was optional, however, and is abandoned in the Koine. The negative in such declarative clauses is οὐ.

Finite clauses can also be used for non-declarative complements. Verbs of effort, such as ἐπιμελοῦμαι ‘take care that’, are complemented by clauses introduced by ὅπως with the future indicative. The negative is μή. By the same sequence-of-mood rule mentioned above the indicative may, in historic sequence, be replaced by the future optative: ἐπιμελοῦμην ὅπως (μή) ταῦτα ποιήσει (ποιήσῃ) ‘I took care that he should (not) do this’. Verbs of fear for the future are followed by μή (negative μή οὐ) with the subjunctive (optionally the optative in historic sequence): φοβοῦμαι μή (οὐ) γένηται ‘I fear it will (not) happen’.

Non-finite complements use either an infinitive or a participle. Declarative infinitive complements follow verbs of thinking and some verbs of saying. Such infinitives express tense as well as aspect – the present infinitive standing for a present or imperfect indicative, future infinitive for future indicative, aorist for aorist, perfect for perfect – and the negative is οὐ. If the subjects of the infinitive and of the matrix verb are co-referential, the subject of the infinitive is left covert and anything in agreement with it is nominative. If the two subjects are not co-referential the subject of the infinitive is overt and accusative. (Note that this is a substantial difference from Latin, where the subject of an accusative and infinitive construction is compulsorily expressed except when the matrix verb is passive.)

Non-declarative infinitives are used after verbs of wishing, commanding, ability etc. Most of these are control structures (subject control in (ἐγώ<sub>i</sub>) βούλομαι [PRO<sub>i</sub> ταῦτα ποιῆν] ‘I want to do this’, object control in κελεύω αὐτόν<sub>j</sub> [PRO<sub>j</sub> ταῦτα ποιῆν] ‘I order him to do this’). Such infinitives express aspect but not tense (the present is used for ongoing or repeated actions, the aorist for simple actions; the perfect is rare, and the future not permitted).

Participles, similarly, may be used in declarative and non-declarative complements. Declarative participial complements are marked for factivity, and as such follow verbs of perception and knowledge. The verb φαίνομαι ‘seem, appear’ selects either a factive participle complement (φαίνομαι ὄν = ‘it is obvious that I am’) or a non-factive infinitive complement (φαίνομαι εἶναι = ‘I appear to be’, a raising structure).

Adjectival and adverbial subordinate clauses follow the same range of strategies. Relative and temporal clauses are introduced by relative pronouns and temporal conjunctions respectively (ὅς ‘who’, ὅτε ‘when’ etc.) and usually have indicatives. Indefiniteness (‘whoever’, ‘whenever’) is expressed by replacing the indicative with the subjunctive along with the modal particle ἄν (optionally in historic sequence by the optative without ἄν). Temporal clauses referring to the future are treated usually as indefinite and behave the same way. Causal clauses introduced by conjunctions such as διότι ‘because’ also have their verb in the indicative; but if the speaker wishes not to vouch for the veracity of the reason expressed, the clause becomes virtually one of reported speech and may, in historic sequence, take the optative.

An important class of finite adverbial clause is that introduced by ἵνα, ὥς or ὅπως with a verb in the subjunctive (optionally the optative in historic sequence), used to express purpose or intention.

Infinitives are rarer in non-complement subordinate clauses than in complements. When the temporal conjunction πρὶν means ‘before’ rather than ‘until’ (so always in Attic after verbs in the affirmative) it governs an infinitive. Result clauses introduced by ὥστε have an indicative if the consequence actually did (or did not, when the negative is οὐ) follow, and an infinitive if the intention was that the consequence should (or should not, μή) follow.

Participles, on the other hand, are abundant and are able to substitute for a wide range of adjectival and adverbial finite clauses, including relative and temporal clauses and expressions of manner, either with or without introductory particles such as ἄτε or ὥς, both ‘because’, καίπερ ‘although’. Where the subject of such a participle is not co-referential with another constituent of the sentence it defaults to the genitive case, a construction known as the genitive absolute. See Ruppel 2013 for a discussion of this construction and its origins.

It will be seen that other than in future remote conditions and in wishes the optative had become an optional substitute for the subjunctive or indicative in certain constructions. It is no surprise, then, that it was lost in the Koine. Other patterns of complementation were also simplified. The complementary participle lost ground to the accusative and infinitive construction after verbs of perception and knowledge; and the declarative accusative and infinitive itself lost ground to finite complements introduced by ὄτι and ὥς. In control structures the infinitive also had a tendency to be replaced by finite subjunctive complements introduced by ἵνα. This in turn lost its force ‘in order that’ and became grammaticalised first as a generic complementiser for subjunctive clauses and then, in a phonetically reduced form να, as a marker of the subjunctive itself.

### Prepositions

The class of words known to traditional grammars as prepositions, denoting such notions as spatial relations and selecting for nominal arguments in particular grammatical cases, originate as independent adverbials. Spatial relations were, at an early stage of the language, encoded solely in case morphology – acc. for ‘motion towards’ (allative sense), gen. for ‘motion away from’ (ablatival sense) and loc., later dat., for ‘place where’ (locative sense). Traces of this system are preserved in Mycenaean, which uses bare forms with dat./loc./instr. morphology to express locative sense (dat.-loc. *Sp<sup>h</sup>agiamp<sup>h</sup>i*, instr. *Sp<sup>h</sup>agiamp<sup>h</sup>i*, both ‘at Sphagianes’), as well as in poetry. Adverbials were used to further specify or nuance the precise relation involved, e.g. παρά, which in one set of uses implies the presence of some person, so παρά τινα ‘into someone’s presence’, παρά τινος ‘away from someone’s presence’, παρά τινι ‘in someone’s presence’.

Where these adverbials occurred immediately before the verb they became regarded as prefixes or preverbs. Their original autonomy is attested by the phenomenon of tmesis in Homer, where what from a classical perspective is a prefix appears to have become separated from its verb. This must reflect a very ancient stage indeed, as there are no examples of tmesis in Mycenaean.

Where the adverbial occurred immediately before a noun, however, it is easy to see how it could become regarded as a head governing a case form, i.e. a preposition. Some, where the sense was appropriate, came to govern all three local cases; others only two (e.g. ἐν + dat. = ‘in (locative)’, + acc. = ‘into (allative)’ – although in this latter sense it was recharacterised in Attic-Ionic and West Greek as ἐν-ς, whence εἰς); and yet others only one (e.g. ἀπό, ‘away from’ with gen. only).

### Word formation

An old pattern of derivation has the *e*-grade in verbs and the *o*-grade in derived nouns. When accented on the root they are often (result of) action nouns, on the suffix often agents: e.g. λέγω ‘speak’, λόγος ‘speech’; τρέφω ‘nourish’, τροφός ‘nurse’. Many have an -ᾶ suffix, e.g. τροφή ‘nourishment’.

There is a rich derivational morphology. Buck (1955a: 315–348), although rather old-fashioned, collects a catalogue of morphs. Of particular note, the PIE agent noun suffix \*-tēr-/tōr- tends in Attic to be replaced by -της (e.g. σωτήρ ‘saviour’, ῥήτωρ ‘speaker’, but Att. κριτής ‘judge’ = dialectal κριτήρ). The -ευσ suffix is very common in occupational terms (ἵππεύς ‘horseman’ etc.). From PIE \*-yo- and \*-yeh<sub>2</sub>, Gr. -ιον forms neuter diminutives (παιδίον ‘small child’) and -ία feminine abstracts (σοφία ‘wisdom’). Action nouns are derived with \*-tis, becoming -σις (mostly abstract, e.g. πίστις ‘trust’) and -ματ- (tending to become concrete, πράγμα ‘thing’ from πράττω ‘make, do’). The PIE adjectival suffix \*-yos forms adjectives of a very general type (πλούσιος ‘wealthy’ from πλοῦτος ‘wealth’), and becomes a standard way of forming patronymics in some dialects. Its *e*-grade \*-eyo- forms adjectives of material (χρῦσεος, Att. χρῦσους ‘gold’). The suffix -ικο-, although as yet unattested in Mycenaean and rare in Homer, has become the most productive formant of general (mostly denominal) adjectives in Attic prose.

Greek also forms compounds extremely freely. An extensive survey is provided by Meißner and Tribulato (2002), to which the reader is referred.

### FURTHER READING

The most recent comprehensive reference works are those of Bakker (2010), who includes chapters on all aspects covered here and many others, and Colvin (2014). Additionally, for the dialects, see Colvin 2007; Buck 1955b contains a useful collection of dialect inscriptions with commentary but is rather dated. For the history of Greek from Great Attic onwards, see Horrocks 2010. On the decipherment of Linear B see Chadwick 1958 and Fox 2013, which expands on Kober’s contribution. Ventris and Chadwick’s seminal *Documents in Mycenaean Greek* (1956; 2d ed. 1972) remains an essential introduction to the Linear B documents and their language; a third edition is at the time of writing in preparation under the editorship of John Killen and Anna Morpurgo Davies. Until then the most current survey of the state of the discipline is provided by the three-volume *Companion* edited by Duhoux and Morpurgo Davies (2008, 2011, 2014). On Cypro-Minoan see Ferrara 2012–13 and Steele 2013; on the Cypriot dialect and its writing system see Egetmeyer 2009. For the pronunciation of Greek see Allen 1987a and Sturtevant 1940, and for the historical developments in phonology Lejeune 1972. On morphology see Chantraine 1991. On the syntax of the verb see Rijksbaron 2002. On nominal formation see Chantraine 1933.

### REFERENCES

- Ahrens, H.L. 1843, *De Graecae Linguae Dialectis*, vol 2., Gottingen
- Allen, W.S. 1959, “Some remarks on the structure of Greek vowel systems,” *Word* 15: 240–251
- Allen, W.S. 1987a, *Vox Graeca: A Guide to the Pronunciation of Classical Greek*, 3rd edn, Cambridge
- Allen, W.S. 1987b, “The development of the Attic vowel system: Conspiracy or catastrophe?,” in: Killen, J.T., Melena, J.L. & Olivier, J.P. (eds.), *Studies in Mycenaean and Classical Greek Presented to John Chadwick*, Salamanca: 21–32
- Andersen, Ø. & Haug, D. 2012, *Relative Chronology in Early Greek epic Poetry*, Cambridge
- Bakker, E.J. 2010, *A Companion to the Ancient Greek Language*, Oxford
- Brixhe, C. 1976, *Le dialecte grec de Pamphylie: Documents et grammaire*, Paris

- Buck, C.D. 1955a, *Comparative grammar of Greek and Latin*
- Buck, C.D. 1955b, *The Greek Dialects: Grammar, Selected Inscriptions, Glossary*, Chicago
- Chadwick, J. 1958, *The decipherment of Linear B*, Cambridge
- Chadwick, J. 1992, "The Thessalian Accent," *Glotta* 70/1/2: 2–14
- Chantraine, P. 1933, *La formation des noms en grec ancien*, Paris
- Chantraine, P. 1991, *Morphologie Historique du grec*, 3rd edn, Paris
- Colvin, S. 2007, *A Historical Greek Reader : Mycenaean to the Koiné*, Oxford
- Colvin, S. 2014, *A Brief History of Ancient Greek*, Chichester
- Dik, H. 1995, *Word Order in Ancient Greek: A Pragmatic Account of Word Order Variation in Herodotus*, Universiteit van Amsterdam, Amsterdam [PhD dissertation]
- Dover, K.J. 1960, *Greek Word Order*, Cambridge
- Duhoux, Y. & Morpurgo Davies, A. (eds.) 2008, *A Companion to Linear B. Mycenaean Texts and their World*, vol 1. BCILL 120, Louvain-la-Neuve
- Duhoux, Y. & Morpurgo Davies, A. (eds.) 2011, *A Companion to Linear B. Mycenaean Texts and their World*, vol. 2. BCILL 127, Louvain-la-Neuve
- Duhoux, Y. & Morpurgo Davies, A. (eds.) 2014, *A Companion to Linear B. Mycenaean Texts and their World*, vol. 3. BCILL 133, Louvain-la-Neuve,
- Egetmeyer, M. 2009, *Le dialecte grec ancien de chypre*, 2 vols, Berlin
- Ferrara, S. 2012–13, *Cypro-Minoan inscriptions*, 2 vols, Oxford
- Fox, M. 2013, *The Riddle of the Labyrinth: The Quest to Crack an Ancient Code and the Uncovering of a Lost Civilisation*, London
- García Ramón, J.L. 1975, *Les Origines postmycéniennes du groupe dialectal éolien. Étude linguistique*, Salamanca
- Horrocks, G. C. 1987, "The Ionian epic tradition: Was there an Aeolic phase in its development?," in: Killen, J.T., Melena, J.L. & Olivier, J.P. (eds.), *Studies in Mycenaean and Classical Greek Presented to John Chadwick*, Salamanca: 269–294
- Horrocks, G.C. 1990, "Clitics in Greek: A diachronic review," in: Rousseau, M. & Panteli, S. (eds.), *Greek Outside Greece II*, Athens: 35–52
- Horrocks, G.C. 1997, "Homer's Dialect," in: Morris, I. & Powell, B. (eds.), *A New Companion to Homer*, Leiden: 193–217
- Horrocks, G.C. 2010, *Greek: A History of the Language and its Speakers*, 2d edn, Oxford
- Jeffery, L.H. 1990, *The Local Scripts of Archaic Greece*, Revised edn, Oxford
- Lejeune, M. 1972, *Phonétique historique du mycénien et du grec ancien*, Paris
- Meißner, T. & Tribulato, O. 2002, "Nominal Composition in Mycenaean Greek," *TPhS* 100/3: 289–330
- Mendéz Dosuna, J. 2012, "Ancient Macedonian as a Greek dialect: A critical survey on recent work," in: Giannakis, G.K. (ed.), *Ancient Macedonia: Language, History, Culture*, Thessaloniki: 133–145
- Morpurgo Davies, A. 1966, "An instrumental-ablative in Mycenaean?," in: Palmer, L. R. & Chadwick, J. (eds.), *Proceedings of the Cambridge Colloquium on Mycenaean Studies*, Cambridge: 191–202
- Morpurgo Davies, A. 1987, "The Greek Notion of Dialect," *Verbum* 10: 7–28
- Parker, H.N. 2008, "The Linguistic Case for the Aiolian Migration Reconsidered," *Hesperia: The Journal of the American School of Classical Studies at Athens* 77/3: 431–464
- Parry, M. 1928a, *Les formules et la métrique d'Homère*, Paris
- Parry, M. 1928b, *L'épithète traditionnelle dans Homère : essai sur un problème de style homérique*, Paris
- Parry, M., Parry, A. & Lord, A.B. 1971, *The Making of Homeric Verse: The Collected Papers of Milman Parry*, Oxford



- Probert, P. 2003, *A New Short Guide to the Accentuation of Ancient Greek*, Bristol
- Rijksbaron, A. 2002, *The Syntax and Semantics of the Verb in Classical Greek: An Introduction*, 3rd edn, Amsterdam
- Ruppel, A. 2013, *Absolute Constructions in Early Indo-European*, Cambridge
- Sihler, A.L. 1995, *New Comparative Grammar of Greek and Latin*, Oxford
- Smyth, H.W. & Messing, G.M. 1956, *Greek Grammar*, Revised edn, Cambridge, MA
- Steele, P.M. 2013, *A Linguistic History of Ancient Cyprus: The Non-Greek Languages, and their Relations with Greek, c. 1600–300 BC*, Cambridge
- Sturtevant, E.H. 1940, *The pronunciation of Greek and Latin*, 2d edn, Philadelphia
- Thompson, R.J.E. 1998, “Instrumentals, datives, locatives and ablatives: the  $-\phi\iota$  case form in Mycenaean and Homer,” *PCPS* 44: 219–250
- Thompson, R.J.E. 2000, “Prepositional usage in Arcado-Cypriot and Mycenaean: a Bronze Age isogloss?,” *Minos* 35: 395–430
- Thompson, R.J.E. 2014, “Mycenaean  $-\pi\iota$  and  $pa-ro$  in the light of TH Uq 434,” in: Bernabé, A. & Luján, E. R. (eds.), *Donum Mycenologicum: Mycenaean Studies in Honour of Francisco Aura Jorro*, Louvain-la-Neuve: 173–187
- Ventris, M.G.F. & Chadwick, J. 1956, *Documents in Mycenaean Greek*, 1st edn, Cambridge
- Ventris, M.G.F. & Chadwick, J. 1972, *Documents in Mycenaean Greek*, 2d edn, Cambridge

# ITALIC

*Rex Wallace*

## INTRODUCTION

As a linguistic term, “Italic” refers most commonly to a group of closely related Indo-European languages spoken throughout the Italian peninsula in the 1st millennium BCE (Map 6.1). The major language in this group, that is to say, the one for which there is a significant documentary corpus, is Latin. It is also the only Italic language – in fact the only language of ancient Italy, apart from Greek – that did not become extinct. The Romance languages, spoken today throughout Western Europe and the Americas, are its modern descendants.

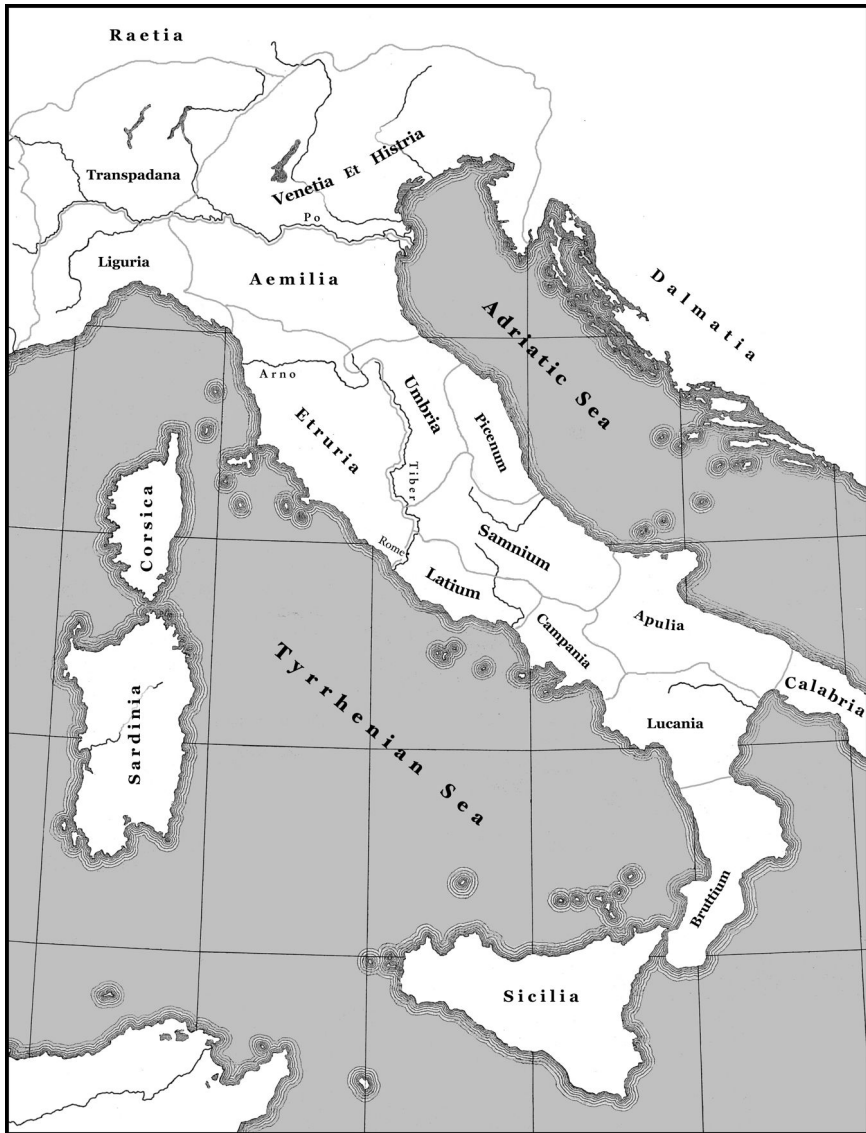
Given the prominence of Latin both linguistically and culturally, and given its importance for Indo-European studies, it is the focus of this chapter. However, data from other members of the Italic branch will be discussed whenever appropriate.

### The languages of ancient Italy

At the beginning of the 1st millennium, Latin was but one of many languages spoken in ancient Italy (see Map 6.2). Celtic-speaking tribes settled in the north, in the region of Lake Maggiore, Lake Como, and the Canton of Ticino. The Veneti inhabited Venetia, from the mouth of the Po River as far as the territory of Istria. The Raeti, who spoke one of the two major non-Indo-European languages in ancient Italy, lived in the sub-Alpine regions around Lake Garda and the Brenner Pass. Speakers of the other major non-Indo-European language, Etruscan, occupied Etruria, but pockets of Etruscan speakers could be found in Latium, Campania, and Aemilia. The Falisci inhabited a small slice of land on the west side of the Tiber River known as the *Ager Faliscus*. Their major center was *Falerii Veteres* (modern Civita Castellana). The Sabellic languages were spoken in an area that extended from Umbria along the spine of the Apennines as far south as Bruttium and Lucania. Messapic-speaking peoples settled Apulia. Dialects of ancient Greek – Ionic, Doric, and Attic – were spoken at numerous colonial outposts in southern Italy.

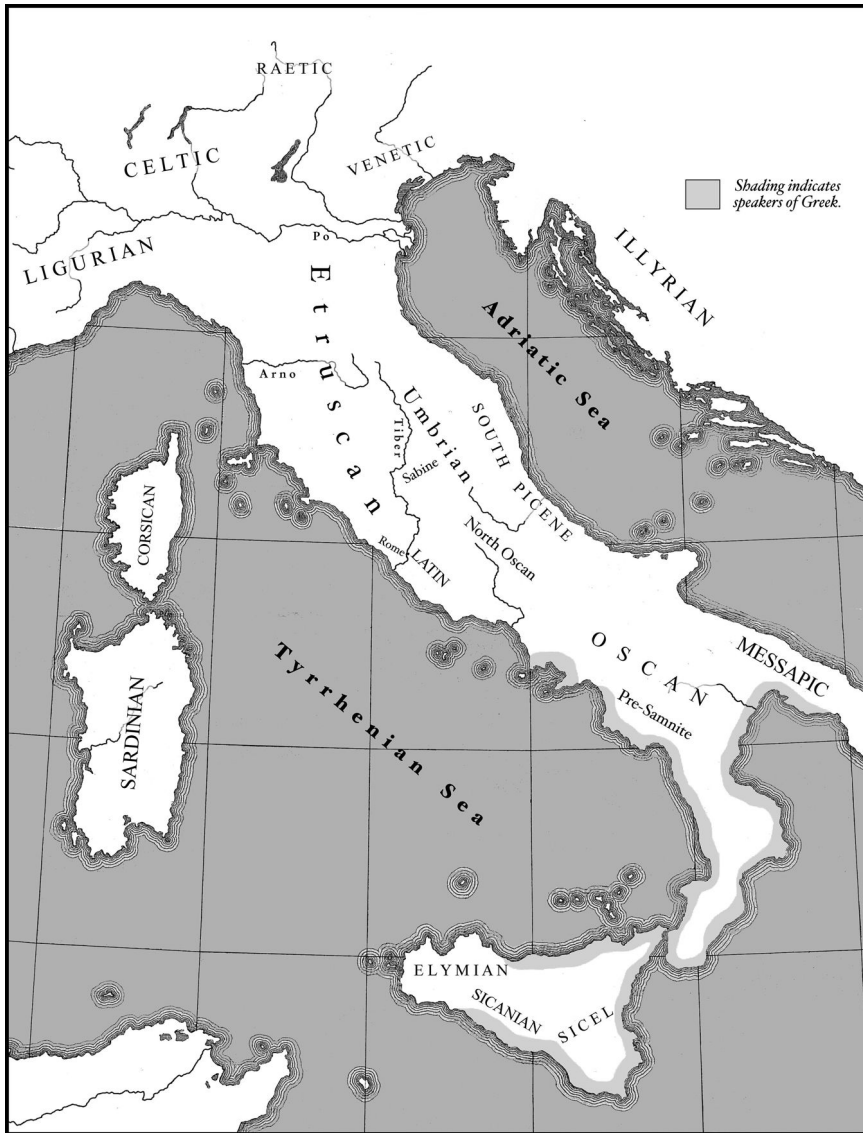
### Italic within Indo-European

Within the Italic branch, Latin was most closely related to Faliscan. The Sabellic languages (once referred to by scholars as Osco-Umbrian) formed the other branch of Italic (Rix 1994). The most important among the Sabellic languages were Oscan, the language of the inhabitants of Samnium and Campania; Umbrian, the language of the inhabitants of ancient Umbria; and South Picene, the language of the people who inhabited the southernmost parts of Picenum along the Adriatic coast. Other members of this branch are known, but only from a handful of inscriptions.



MAP 6.1 THE REGIONS OF ANCIENT ITALY

Venetic may have been Italic as well, perhaps having its own branch. The amount of evidence that can be used to determine the position of Venetic is limited and difficult to evaluate. Phonological developments, such as the change of Proto-Indo-European (PIE) aspirates to fricatives in word-initial position and to voiced fricatives or stops in medial position, point to an Italic connection, but the morphological evidence, such as it is, does not contribute to a decision one way or the other. The best approach is to leave the question open and await the acquisition of additional evidence.



MAP 6.2 THE LANGUAGES OF PRE-ROMAN ITALY

### Italo-Celtic

Correspondences, both phonological and morphological, are shared by members of the Celtic and Italic branches: the thematic genitive singular in  $-ī$  (Old Irish *MAQQI* ‘son’, Latin *uirī* ‘man’, Faliscan *titi* /*titi:*/ ‘Titus’ [personal name]); the  $\bar{a}$ -subjunctive (Old Irish *·bera*, Latin *ferat* ‘carry’, Oscan *putiād* ‘be able’); the superlative suffix  $*(i)s̄m̄mo-$  (Old Irish *tressam* ‘strongest’, Latin *maximus* ‘greatest’); and the development of  $*CR̄HC$  to

*CrāC* (Old Irish *grán*, Latin *grānum* ‘grain’ < \*ǵrHnom). The probative value of some correspondences, e.g., the genitive in *-ī*, which is found also in Messapic, is not easy to assess (Clackson & Horrocks 2007: 31–34, Weiss 2009: 465–466, Fortson 2010: 276–277). As a result, scholars’ opinion on the question of an Italo-Celtic unity remains mixed.

### Origins of Latin

At the beginning of the historical period Latin was spoken by the inhabitants of *Latium* (see Maps 6.1 and 6.2). This region was less extensive than the modern province of Lazio. It was bounded to the north by the Tiber River, to the east by the Sacco and Liri river valleys, to the south by the Garigliano River, and to the west by the Tyrrhenian Sea.

Archaeological evidence points to the development of a regionally distinct material culture, the so-called Latial culture, in *Latium* during the late Bronze Age, ca. the 13th–11th centuries BCE (Holloway 1994: 13–14). Evidence for habitation at Rome in the earliest stages of the Iron Age, ca. the 10th and 9th centuries BCE, is found on the Palatine and Capitoline Hills, and in the Forum area. The evidence for habitation is slightly later at other sites in *Latium*, for example, at Ardea, Satricum, Gabii, Tibur, and Praeneste. Although it cannot be proved, it is likely that the people who produced the artifacts attributed to “Latial” culture in the late Bronze and early Iron Ages were speakers of Latin.

Speakers of the Italic languages probably entered the Italian peninsula from the north-east by crossing over the Alps. When this happened cannot be determined, but the idea that Italic-speaking peoples had begun to settle in Italy before the middle of the 2nd millennium BCE, if not earlier, is not likely to be wrong. The linguistic geography of pre-Roman Italy leads us to surmise that Italic speakers spread southward from northeastern Italy along the spine of the Apennines and fanned out into the more hospitable valleys and coastal plains of central and southern Italy.



**FIGURE 6.1** *FIBULA PRAENESTINA* (CIL 1<sup>2</sup>.3). Reproduced by permission of the Center for Epigraphical and Palaeographical Studies, the Ohio State University

## The spread of Latin

Apart from Greek, Latin was the only language of ancient Italy that did not become extinct. Speakers of other languages succumbed to the military, political, and cultural domination of Latin. The Romans captured the Faliscan capital of *Falerii Veteres* in 241 BCE and forcibly removed all of its inhabitants to a less defensible position. Faliscan speakers switched to Latin during the 2nd century BCE. Etruscan and the Sabellic languages, Oscan and Umbrian, were still spoken at the beginning of the imperial period, but it is unlikely that native speakers of these languages survived beyond the final decades of the 1st century CE. Even the languages – both Indo-European and non-Indo-European – spoken at the geographic extremes of the peninsula disappeared soon after Roman legions and Roman administration gained a foothold in their territories. Venetic, Raetic, and Celtic, languages spoken in the Cisalpine territories of northern Italy, did not survive much beyond the reign of Augustus. Messapic inscriptions are not attested after the 1st century BCE. Ancient Greek was the only language spoken on the peninsula in antiquity that was not supplanted by Latin. It survived primarily because of its prestige value for Romans and because its speakers maintained a vibrant, though often contentious, relationship with the homeland.

In the course of a few centuries, then, Latin went from being the language of *Latium* to the language of a vast empire that included Europe and parts of Africa and Asia. The paths of diffusion of the language coincided with Roman military conquests. By the end of the war against Hannibal and the Carthaginians, Rome had annexed the islands of Sicilia, Sardinia, and Corsica, and set up outposts in southern Spain. Annexation of Illyricum (which was situated along the Dalmatian coast), of northern Africa, and of Greece followed shortly thereafter. The final pieces of the Roman imperial conquest, Britain and Dacia, were in place by the end of the 1st century CE.

Although the Romans never developed an official language policy, conquered peoples eventually abandoned their native languages for Latin. The Roman ruling classes, who were responsible for civil administration, finances, and military matters in the colonies and conquered territories, left most institutions and religious practices in the hands of the natives. But the fact that Latin was perceived as the language of prestige, and thus the language of economic, political, and military advancement, eventually made it the language of choice for most speakers. The adoption of Latin took place via periods of bilingualism during which the natives spoke their own languages at home in private with family and friends, and Latin in the public arena. Ultimately, great numbers of speakers, within Italy and outside of Italy, gave up their native languages for Latin. Pre-existing languages generally disappeared without a trace, although some scholars have argued that it is possible to detect features of the substratum languages by means of changes appearing in the Romance languages.

Ironically, prestige is the reason why Latin did not succeed in replacing ancient Greek as the language of everyday use in the eastern part of the empire. For the Roman elite, the Greek language enjoyed an especially elevated status. Members of the Roman aristocracy, when they could afford to do it, sent sons to Athens for study as part of a rigorous program of education.

## Latin and the Romance languages

Latin is the linguistic parent of the Romance language group and as such is the progenitor of the modern Romance languages. French, Italian, Portuguese, Rumanian, Spanish, and

the remaining Romance varieties are the modern continuants of regional Latin varieties that were spoken by traders, commercial entrepreneurs, soldiers, administrators, and their families at the colonial settlements in foreign territories incorporated into the Roman Empire.

### Medieval Latin and later Latin

Latin did not cease to be written even as regional varieties of the spoken language diverged and ultimately emerged as Romance languages, nor did it cease to be spoken in the academic arena. Latin was adopted as the *lingua franca* of scholarly discourse during the medieval period, and it survived as such well into the formative period of the Renaissance, finally giving way to treatises and tracts written in the native tongues of their authors. With the rise of linguistic nationalism, Latin was restricted more and more to the sphere of religious discourse and learned publications. Indeed, the former is the primary venue in which Latin survives today. The Vatican continues to publish its encyclicals in Latin.

### Varieties of Latin

Although it is impossible to say when Latin and the other Italic languages began to diverge, we can pinpoint the date of the earliest inscriptions. They go back to the middle of the 7th century BCE. And, although we cannot say precisely when Latin ceased to be Latin and the Romance languages became recognizably such, we can point to the earliest text attested in a Romance language, e.g., the Strasbourg Oath in Old French, which dates to 842 CE, and assume that the “transition” from Latin to Romance happened somewhere between the 7th and 9th centuries CE. If we accept this as the transition period between Latin and the Romance languages, then the recorded history of the Latin language ranges over some 1,250 years.

Because the history of Latin covers a vast expanse of time, it is possible to recognize changes in the language at all levels of linguistic structure over the course of its development from the language of the earliest inhabitants of *Latium* to the language of the inhabitants of the military outpost at Vindolanda in Great Britain. Indeed, since the chronological period for Latin is so great, it is convenient to recognize stages in the historical development of the language. Even though these chronological divisions are arbitrary, they serve as guideposts that facilitate discussion. Most historians of the Latin language accept some version of the following schema:

- Very Old Latin (VOLat.), ca. 650–400 BCE
- Old Latin (OLat.), ca. 400–100 BCE
- Classical Latin (Class. Lat.), ca. 100 BCE–200 CE
- Late Latin (Late Lat.), ca. 200–600 CE

Looking across the chronological divisions of the language, it is possible to identify differences in pronunciation, in the formation of words, and in sentence construction. For example, the dative singular of *o*-stem nouns and adjectives ended in a long diphthong in Very Old Latin, e.g., *DUENOI* /dwenɔːj/ ‘good (man)’, m. dat. sg. By the Old Latin period, however, this long diphthong had changed to a long vowel; other changes transformed the initial syllable of *DUENOI* from \*dwe- to *bo-*, yielding Classical Latin *bono* /bɔnoː/. In Old Latin the syntax of *cum*-clauses, regardless of their function, called for a verb in the

indicative mood. By the Classical Latin period, however, the mood of the verb in clauses with causal and concessive function was the subjunctive. In Late Latin the preposition *de* ‘from’ takes on the functions of the genitive case; that is to say, it can be used to indicate possession, part for whole, and the “objective genitive”.

While it is important to recognize chronological differences within the history of Latin, it is equally important to recognize differences based on geography, social class, and style or register. As an example of regional variation we point to the dative singular of *ā*-stem nouns. In the Latin of the 3rd century BCE, speakers of some Latin dialects outside of Rome pronounced the dative singular ending with a long *ā*, rather than with a diphthong, which was the standard in Rome, e.g., FORTUNA /fōrtu:na:/ ‘luck’ (Praeneste), but Roman Latin *mēnsae* /me:nsae/ ‘house’. In the Latin of Pompeii – but only in graffiti and thus perhaps only in the speech of sub-elite social orders – syllable-final short *i* was syncopated when standing between /w/ and /t/, e.g., ABERAUT ‘has lost’, cf. CL *aberrāuit*. This development is not found in elite Latin texts of the 1st century CE, which suggests that it was restricted, at least at its inception, to sub-elite Latin dialects. The language variety employed for literary composition, which is commonly referred to as Classical Latin, was a *Kunstsprache*. It drew on features from a number of other varieties of Latin, for example, Old Latin, and the language of law and religion. The writers of this variety also appropriated or elaborated on linguistic and stylistic features borrowed from ancient Greek authors (Clackson & Horrocks 2007: 183–228).

The Latin language, then, as we would expect of any language, covered a spectrum of varieties, written and spoken, geographic and social, formal and informal. Some scholars adopt a more severe approach, viewing the differences between elite and sub-elite varieties of Latin in terms of *diglossia*, in which two co-existing linguistic systems are recognized, one for the language of literature and official communication, the other for the language of everyday speech. The idea is that these two systems were already divergent by the end of the Roman Republic, and that they continued to diverge over the course of the history of the empire and beyond. The evidence does not corroborate this position.

### Documentary evidence

Latin has come down to us through the medium of written texts. There are two primary documentary sources. Arguably, the most important source comes in the form of manuscripts of literary texts. Unfortunately, few Latin manuscripts have survived from antiquity. Most are late, dating as they do from the 10th to the 15th centuries CE. Thus, the texts of Roman authors are attested at least a thousand years after they were originally composed. For this reason the orthography of the manuscripts, particularly the manuscripts of the Old Latin poets, may not be as reliable as one would wish it to be. Editorial changes made by later copyists, who sometimes altered Old Latin spellings to conform to Later Latin models, have made their way into the manuscript tradition.

Roman literary compositions are the nucleus of this form of documentation. The literary productions of the most distinguished Roman poets and prose writers, e.g., Vergil, Catullus, Ovid, and Horace for poetry, and Caesar, Cicero, Livy, and Tacitus for prose, are known to us with few exceptions from manuscripts. But so also are the works of authors such as Columella and Vitruvius, both of whom composed important technical treatises, the former on agriculture, the latter on architecture.

The second important source of documentation is epigraphic. Epigraphic documentation comes in the form of inscriptions, graffiti, and dipinti (painted texts) that have been discovered over the length and breadth of Roman territory. For the most part, the



inscriptions that have survived were incised or chiseled on non-perishable material – ceramic, stone, and metal.

Some inscriptions were produced in order to stand the test of time: epitaphs preserving the memory of the deceased, texts of laws and decrees incised on bronze and earmarked for public display, texts commemorating generous contributions by patrons to their communities, dedications intended to accompany an offering to a deity, and the “signatures” of ceramic makers or their owners. Other texts survived on account of favorable circumstances. The writing on the wooden tablets of the Pompeian banker *Lucius Caecilius Iucundus* survived because of historical accident, namely, the eruption of Mt. Vesuvius in 79 CE. Lead tablets incised with curses against competitors or against rivals in love survived because, in order to be efficacious, they had to be devoted to underworld demons and were therefore hidden away in tombs, buried underground, or tossed into wells. The leaf-tablets from Vindolanda survived because they were abandoned in drainage ditches that were subsequently covered with clay, which deprived the bacteria that would have destroyed them of the oxygen necessary to multiply.

### ALPHABET, ORTHOGRAPHY, AND PHONOLOGY

The Latin alphabet was borrowed from the Etruscans, probably from the inhabitants of Caere, a settlement located across the Tiber River a short distance from Rome. The Latin *abecedarium* cited in (1) is the result of several alphabetic reforms that took place between the 7th and 3rd centuries BCE.

(1) A B C D E F G H I K L M N O P Q R S T U X

The letter <G> was added to the alphabet Latin speakers inherited from the Etruscans. It first appears in the 2nd half of the 3rd century BCE and, oddly enough, occupied the position once held by <Z>, which seems never to have been used to write Latin. The letter <F>, which originally had the value /w/, is the result of the simplification of the digraph <FH>, an older spelling for the sound /f/. In the early imperial period the Greek letters <Y> and <Z> were added to the script to spell sounds found in Greek loan words. They were placed at the end of the *abecedarium*. An attempt to add additional consonantal signs to the alphabet was made by the emperor Claudius, and though two of his signs were used in inscriptions, they disappeared from the writing system soon after his death.

For the most part, the consonantal letters stood in a one-to-one relationship to the consonantal phonemes in the sound system, but in a few instances the spelling of sounds was less optimal. The letters <C> and <K> stood for the voiceless velar /k/, e.g., *capiō* ‘seize’ and *kalendae* ‘Kalends’. Occasionally, <Q> spelled /k/, especially if the following vowel sign was <U>, e.g., *PEQUINIA* ‘money’. This three-way spelling of /k/ is the remains of an orthographic rule inherited from the Etruscans. The digraph <QU> spelled a voiceless labiovelar; <GU> probably spelled the voiced counterpart of <QU> (on which see below). Finally, the letter <X> stood for the cluster /ks/, e.g., *arx* ‘citadel’ /arks/.

Vowel length was rarely represented orthographically. A few attempts at indicating vowel length, for example, writing vowels double (PAASTORES ‘shepherds’), marking length with a diacritic (MURUM ‘wall’), and employing etymological spellings such as *ei* for /i:/ ([U]EIJAM ‘living’), were used sporadically, but they were not adopted as the standard. Thus, in most instances, a single vowel sign stood for both a long vowel and its short counterpart, e.g., <A> = /a:/ and /a/. The glides /j/ and /w/ were also spelled with vowel signs, e.g., *iam* ‘now’ = /jam/ and *uir* ‘man’ = /wir/. In the Classical Latin spelling system, the signs <I> and <U> represented three phonemes.

Metrical evidence indicates that words like *peior* ‘worse’, *maior* ‘more’, and *aiō* ‘say’ had a heavy first syllable. When standing between vowels, the letter spelled a double or long consonant /majjor/.

In the oldest Latin inscriptions the direction of writing was variable. Some inscriptions were written in left-to-right direction, others in right-to-left direction. A few inscriptions were incised in the boustrophedon style. By the 4th century BCE, however, left-to-right direction was the standard. Punctuation in texts was minimal. Very Old Latin inscriptions were often written *scriptio continua*. Later, punctuation points regularly separated words, but sentence-signs and paragraph-signs were rare, even in long inscriptions.

## Phonology

The phoneme inventory of Latin is displayed in Table 6.1.

The voiced labiovelar, which was restricted to the environment after nasals (*ninguit* ‘it’s snowing’), is the only phoneme whose status is controversial. It is listed as part of the phonological inventory based on phonotactic and morphophonemic evidence. The treatment of *qu* and *gu* as clusters /kw/ and /gw/ does not fit with the metrical evidence, which points to a unitary phoneme, at least as far as *qu* is concerned. Interestingly, the metrical evidence also suggests that the cluster /kw/, as in *equus* ‘horse’ < \*ekwos, fell together with the original labiovelar.

The velar nasal [ŋ] is sometimes considered a separate phoneme based on the evidence of minimal triplets: *amnī* ‘river’, *annī* ‘year’, and *agnī* ‘lamb’ [aŋni:]. However, it is possible to treat [ŋ] as an allophone of the phonemes /n/ and /g/. Following this analysis, the velar nasal in *longus* [lɔŋɡos] is the result of the assimilation of /n/, a dental nasal, to the velar position of the following stop. The velar stops /k/ and /g/ were pronounced as the corresponding nasal when standing before a dental nasal. Thus, *ignis* ‘fire’ was pronounced [iŋnis] and *dignus* ‘worthy’ as [diŋnos].

The glottal fricative /h/ was restricted in large part to word-initial position, and it was a feature of elite Latin. But even in the elite variety of the language, word-initial *h* behaved as an outlier. It did not “make position” in metrical texts, and members of the educated classes, as is clear from poem 84 of Catullus, could be guilty of hypercorrect pronunciations. Although *h* was written in medial position in some words, it is doubtful that it was pronounced. *h* did not prevent vowel contraction, e.g., *mī* for *mihi* ‘to me’, and it did not block rhotacism of *-s*, e.g., *diribeō* for \*dishabeō.

TABLE 6.1 THE PHONEMES OF CLASSICAL LATIN

Consonants										
stops										
	p	b	t	d	k	g		k <sup>w</sup>	g <sup>w</sup>	
fricatives										
	f		s						h	
nasals										
	m		n							
lateral										
		l								
trill										
			r							
glides										
				j				w		
Vowels										
diphthongs										
ī	i				o	ū		aē	ej	uj
	ē	ε		ɔ	ō			aw	ew	oē
		ā	a						ow	

The vowels in the system were distinguished by length. The short members were lax and more central than their long counterparts. The number of diphthongs in the system was for all intents and purposes restricted to /ae/ and /aw/. /oe/ was present in a handful of words, e.g., *poena* /poena/ ‘punishment’, perhaps under Greek influence, cf. ποινή. The diphthongs /ej/, /uj/, etc., were the result of contractions and were restricted to a small number of words, e.g., *deinde* ‘then’, *cui* ‘to whom’.

In Classical Latin, epigraphic and metrical evidence suggest that word-final *-Vm* surfaced as a long nasalized vowel. In Old Latin inscriptions of the 3rd century BCE, final *-m* was often omitted in writing, e.g., OL OPTUMO ‘best’ = CL *optimum*, a spelling that makes sense if the nasal was lost and the preceding vowel was nasalized. The etymologically appropriate spelling was restored in the 2nd century BCE and became standard in literary Latin and bureaucratic prose. In poetry, words ending in *-Vm* were subject to elision if the following word began with a vowel just as if they ended in a long vowel. And words ending in *-Vm* ‘made position’, that is to say were metrically heavy, if followed by a word beginning with a consonant. Nasalized vowels were also found before *-ns*, both primary and secondary, and *-nf-* (nasal + fricative), e.g., *infāns* ‘child’, *mēnsa* ‘table’, *cōnsul* ‘consul’. Note the common abbreviation *cos.* for *cōnsul*.

## Diachronic developments

From a diachronic perspective the most distinctive phonological development, and one that set Italic apart from its sisters, was the treatment of the PIE voiced aspirates (Stuart-Smith 2004).

In word-initial position, the aspirates changed to voiceless fricatives. The labial and dental aspirates merged as /f/; the velar aspirate, reflecting an earlier merger of PIE \*g<sup>h</sup> and \*g<sup>h</sup>, became /h/, presumably via an intermediate stage as \*x. Comparative data are assembled in (2).

### (2) PIE aspirates in initial position

\*b<sup>h</sup>: Latin *frāter* ‘brother’, m. nom. sg., Oscan **fratrūm**, m. gen. pl., Umbrian **frater** ‘brothers’ (in a religious fraternity), m. nom. pl. < \*b<sup>h</sup>reh<sub>2</sub>ter-; Latin *ferō* ‘I carry’, 1 sg. pres. act., Marrucian *ferenter* ‘they are carried’, 3 pl. pres. pass., Umbrian **fertu** ‘he should bring’, 3 sg. imp. act. < \*b<sup>h</sup>er-; Latin *far* ‘grain’, n. nom. sg., Faliscan **far**, Oscan **far**, Umbrian *far* < \*b<sup>h</sup>ars

\*d<sup>h</sup>: Latin *faciō* ‘I make’, 1 sg. pres. act., Oscan **fakiiad** ‘let him sacrifice’ 3 sg. pres. subj., Umbrian **façia** ‘he should sacrifice’, 3 sg. pres. subj., Volscian *façia*, 3 sg. pres. subj., Venetic **vhağsto** ‘he made’, 3 sg. perf. < \*d<sup>h</sup>h<sub>1</sub>-k-, cf. Latin *fēcit*, 3 sg. perf., based on full grade \*d<sup>h</sup>eh<sub>1</sub>-k-; Latin *fānum* ‘shrine’, n. nom. sg. < \*d<sup>h</sup>h<sub>1</sub>s-no-, Oscan **fīsnū** ‘temple, f. nom. sg., Umbrian **fesnaf(e)** ‘to the temple’, f. acc. pl. < \*d<sup>h</sup>eh<sub>1</sub>s-neh<sub>2</sub>-

\*g<sup>h</sup>/g<sup>h</sup>: Latin *hortus* ‘garden’, m. nom. sg., Oscan **hūrz** ‘grove’ (as sacred area), m. nom. sg. < \*g<sup>h</sup>ortos ‘enclosed area’; Latin *horitur* ‘incites’ 3 sg. deponent < \*g<sup>h</sup>ṛ-ye/o-, Oscan **heriiaad** ‘he should wish’, 3 sg. subj. act., Umbrian **heriest** ‘he will wish’, 3 sg. fut. act. < \*g<sup>h</sup>er-ye/o-

\*g<sup>wh</sup>: Latin *formus* ‘warm’, m. nom. sg. < \*g<sup>wh</sup>ormos; *de-fen-dō* ‘I ward off’, 1 sg. pres. act. < \*g<sup>wh</sup>ende/o- (root \*g<sup>wh</sup>en- ‘destroy’) [no Sabellic examples]

In medial position the lines of development were more complex, and they served to distinguish Latin from the rest of the Italic group. The labial and dental aspirates developed

into voiced fricatives [β] and [ð], which may be the stage attested by Venetic, if **lo.u.ze-roφo.s.** ‘children’, dat. pl., is phonetically [lowðeroβos]. Medial [β] and [ð] merged as [β] in Faliscan and in the Sabellic languages. In Rome and environs, [β] and [ð] survived and eventually became the corresponding voiced stops /b/ and /d/. The velars (PIE palatals and velars), on the other hand, developed to /h/ in all Italic varieties, perhaps via a voiced velar fricative. The aspirated labiovelar moved along a different trajectory in Latin and the Sabellic languages: \*g<sup>wh</sup> became *u* /w/ in Latin but *f* /β/ in Sabellic. Examples of the development of aspirates in medial position are in (3).

### (3) PIE aspirates in medial position

- \*b<sup>h</sup>: Latin *albus* ‘white’, m. nom. sg., Umbrian *alfir*, n. abl. pl. < \*(h<sub>1</sub>)alb<sup>h</sup>os, cf. Oscan **alafaternum** ‘inhabitants of Nuceria Alfaterna’, m. gen. pl.; Latin *tibi* ‘to you’, dat. sg., Oscan **t(e)feí**, dat. sg., Umbrian **tefe**, dat. sg., South Picene **tefeí**, dat. sg. < \*teb<sup>h</sup>ey
- \*d<sup>h</sup>: Latin *medius* ‘middle’, m. nom. sg., South Picene **meñin** loc. sg. + postposition -en, Oscan **meñai**, f. loc. sg. < \*med<sup>h</sup>yo-; Latin *ruber* ‘red’, m. nom. sg., Umbrian **rufru** < \*h<sub>1</sub>rud<sup>h</sup>ro-
- \*ǵ<sup>h</sup>/ǵ<sup>h</sup>: Latin *uehō* ‘I transport’, 1 sg. pres. act., Umbrian **ařveitu**, ‘let him bring’, 3 sg. pres. imp. < \*weǵ<sup>h</sup>e/o-; Oscan **feihúss** ‘walls’, m. acc. pl. < \*d<sup>h</sup>eyǵ<sup>h</sup>-, cf. Latin *fiŋgō* ‘I fashion, shape’, 1 sg. pres. act.; Latin *mihi* ‘to me’, dat. sg., Umbrian **mehe**, dat. sg. < \*meǵ<sup>h</sup>ey
- \*g<sup>wh</sup>: Latin *foueō* ‘I heat up’, 1 sg. pres. act. < \*d<sup>h</sup>og<sup>wh</sup>eye/o-; Umbrian **vufu** ‘votive’, m. acc. sg. < \*h<sub>1</sub>wog<sup>wh</sup>-ro-

The development of the PIE labiovelars, \*k<sup>w</sup> and \*g<sup>w</sup>, is another change that distinguished the Sabellic languages from Latin and Faliscan (and from Venetic too). \*k<sup>w</sup> survived in Latin and Faliscan in most environments, e.g., Latin *-que* ‘and’, Faliscan **-cue**, cf. Venetic **-kve** < \*-k<sup>w</sup>e. \*g<sup>w</sup> survived too, but only when it followed a nasal, e.g., Latin *unguen* ‘salve’, n. nom. sg. < \*h<sub>3</sub>eng<sup>w</sup>- ‘anoint’; in other environments, it changed to *u* /w/, e.g., Latin *ueniō* ‘come’, 1 sg. pres. act. < \*g<sup>w</sup>m̄-ye/o- ‘sets out’. In contrast, \*k<sup>w</sup> and \*g<sup>w</sup> merged with the labial stops in Sabellic, e.g., Oscan **pis** ‘who’, m. nom. sg. < \*k<sup>w</sup>is; Oscan **benust** ‘will have come’, 3 sg. fut. perf. < \*g<sup>w</sup>em- (the dental nasal by analogy with the present). For the development of the aspirated labiovelar, see above.

In Latin intervocalic \*s changed to *r*, e.g., Latin *generis* ‘family’, n. gen. sg. < \*geneses. This change, generally referred to as rhotacism, took place also in Umbrian, in final as well as medial position, e.g., Umbrian **farariur** ‘of the grain’, m. nom. pl. < \*-āsīyōs. Oscan seems to have preserved the intermediate stage of development. Medial \*s was voiced but not rhotacized, e.g., *egmazum* ‘property’, f. gen. pl.

The liquids and nasals developed for the most part unchanged. Syllabic liquids developed to oR, e.g., *mors* ‘death’, f. nom. sg. < \*m̄rtis; syllabic nasals developed to eN, e.g., Latin *decem* ‘ten’ < \*dek̄m̄, cf. Umbrian *desen-duf* ‘twelve’ (‘ten-two’). In Sabellic, syllabic nasals in word-initial syllables appear as aN, e.g., **tanginud** ‘decision’, f. abl. sg. This may be a secondary development, perhaps under accent.

The PIE laryngeals were lost in pre- and post-consonantal positions in Italic. If a laryngeal followed a syllabic resonant, the result was Rā, e.g., Latin *grāta*-, Oscan **bratas** ‘favor’, f. gen. sg. < \*g<sup>w</sup>ṛH-teh<sub>2</sub>; Latin *nātus* ‘son’, m. nom. sg., Paelignian *cnatois*, m. dat. pl. < \*ǵñh<sub>1</sub>-to-. Laryngeals standing between consonants were vocalized to *a*, e.g., Latin *datus* ‘given’, nom. m. sg. < \*dh<sub>3</sub>-to-, *animus* ‘spirit’, nom. m. sg. < \*h<sub>2</sub>enh<sub>1</sub>-mo- (medial

-i- here by “vowel weakening”, on which see below, under “Accent”), cf. Oscan **anams** ‘breath’. Original \*VHC sequences yielded long vowels via loss of the laryngeal and compensatory lengthening of the preceding vowel, e.g., Latin *fēriae* ‘holidays’, f. nom. pl., Oscan **fiisnam** ‘sanctuary’, f. acc. sg. Both nouns were derived from the root \*d<sup>h</sup>eh<sub>1</sub>s- ‘divinity’.

In word-initial syllables, PIE short and long vowels survived in Latin in most environments. Long vowels, regardless of source, were shortened in word-final syllables closed by any consonant but -s, e.g., *portāt* ‘he carries’ > *portat*, *portōr* ‘I’m carried’ > *portor*, etc., but *portās* ‘you carry’. These changes led to morphophonemic alternations in Latin verb paradigms (see Table 6.10).

Apart from \*ay and \*aw, the PIE diphthongs developed into long vowels in all positions in Latin, e.g., \*ey > *ī* (\*deykō > *dīcō* ‘say’); \*oy > *ū* (\*oynos > *ūnus* ‘one’); \*ew, \*ow > *ū* (\*dewkō > *dūcō* ‘lead’).

### Accent

At some point in the prehistory of Latin, the pitch-accent system of PIE was replaced by a system in which the vowel in the initial syllable of a word carried a stress accent.

The effects of this stress accent are abundantly evident in Latin. Short vowels were subject to phonetic processes known collectively as “vowel weakening”. Quality distinctions were eliminated in non-initial syllables. In open syllables short vowels changed to *i*, e.g., *faciō* ‘make’ beside *efficiō* ‘construct’. In closed syllables, original *a* changed to *e*, and original *o* to *u*; original *e* and *i* remained, e.g., *factus* ‘made’ beside *effectus* ‘constructed’, *euntis* ‘going’, m. gen. sg. < \*eyontes. Diphthongs in non-initial syllables, which did not otherwise become long vowels (see above, “Phonology”), were also affected by “weakening”. \*ay and \*aw ultimately changed to *ī* and *ū*, e.g., *caedō* ‘I cut’ beside *incidō* ‘I cut into’, and *claudō* ‘I shut’ beside *concludō* ‘I confine’.

Over the course of the history of Latin, unstressed vowels in open medial syllables, usually, but not exclusively, in the environment of sonorants or fricatives, were subject to syncope processes. Examples follow: \*deksiteros ‘right’ > *dexter*; \*uirotūtis ‘manliness’ > *uirtūs*; and \*kedate ‘give here’ > *cette*.

By the 3rd century BCE, the Latin “law of the penult” had replaced the word-initial accent. Penultimate syllables, if heavy, were accented. If the penultimate syllable was light, then the accent shifted to the antepenultimate syllable, if there was one, e.g., *affēctus*, *portāmus*, *dūkimus*, but *dūcit*, *ánimum*, *fāciō*, etc.

The Sabellic languages also show the effects of a word-initial accent. Several rounds of syncope eliminated short vowels in medial and final syllables. In Proto-Sabellic, short vowels in word-final syllables were lost before -s, e.g., Oscan **hurz** ‘grove’, m. nom. sg. < \*g<sup>h</sup>ortos, Oscan **Pakis** ‘Pacius’, m. nom. sg. < \*pakiyos, Oscan **humuns** ‘men’, m. nom. pl. < \*homones. Subsequently, short vowels in medial syllables were syncopeated, e.g., Oscan **actud** ‘let him do’, 3 sg. act. imp. < \*agetōd, Oscan *factud* ‘let him make’, 3 sg. act. imp. < \*fakitōd.

In the Sabellic languages the accent may have remained on the initial syllable of words. The double spelling of vowels in Oscan is found with few exceptions in word-initial syllables, and this points to the continuation of vowel length in this position, presumably under accent.

Whether or not the change from a pitch accent to a stress accent is to be assigned to the Proto-Italic period cannot be determined. Some scholars prefer to see it as a later development that originated among languages spoken in central Italy, which then spread

out from the innovative language via contact. There is something to be said for this view. In the non-Indo-European language Etruscan, vowels in medial syllables were synco-pated, which suggests that the vowel in the initial syllable was stressed, e.g., Old Etruscan **túruce** ‘dedicated’ > New Etruscan **túrce**.

## NOMINAL AND PRONOMINAL MORPHOLOGY

Nominal forms were inflected for the categories of gender (m., f., and n.), number (sg. and pl.), and case (nom., voc., acc., gen., dat., and abl.).

The PIE case system was reduced from eight to seven in Proto-Italic; the ablative and instrumental merged as the ablative. The number of cases was further reduced in Latin by the merger of the locative and the ablative, although the locative singular endings *-ae* and *-ī* are attested in a few nouns and in the names of towns and cities, e.g., *humī* ‘on the ground’, *domī* ‘at home’, *Romae* ‘in Rome’, etc. In Oscan, Umbrian, and the other Sabellic languages, the locative case remained a more vibrant part of the system, e.g., Oscan **eíseí tereí** ‘in this land’. The locative appears to have survived in Venetic too, to judge from the noun phrase **dekomei diei** ‘on the 10th day’, cf. Latin *dīē septimī* ‘on the 7th day’.

The dual was lost as a morphological category in Italic, but a few forms, e.g., Latin *ambō* ‘both’ and *duo* ‘two’, preserve the dual nominative/accusative ending. Remnants of the PIE collective survived in Latin *o*-stem plurals, such as *loca* ‘places’ beside the plural *locī*, and perhaps indirectly in nouns such as *pīla* ‘ball (of hair)’ beside *pīlus* ‘a hair’.

Nominal stems had the grammatical feature of gender. Certain stem formations were associated with certain genders. For example, *ā*-stems (*\*-eh<sub>2</sub>*-stems) were overwhelmingly feminine; *o*-stems were masculine and neuter; *-men*-stems were neuter, and so forth. Sometimes nouns of the same formation could be either masculine or feminine, depending on the gender of the referent, e.g., *pater* ‘father’, *frāter* ‘brother’, and *māter* ‘mother’.

### The case endings

In some noun classes, perhaps within late PIE, contraction of the stem-final vowel and the vowel of the inflectional ending led to forms where it was not easy to distract stem and ending. Consider the dative singular of *o*-stems in which the ending *-ō* (VOL *-oi*) was the result of the contraction of *\*-o-ey* > *\*-ōy* and then the loss of diphthong-final *y*, e.g., VOL *DUENOI* ‘good [man]’, CL *bonō* ‘good’. The same is true also for many cases in the *ā*-stem (< *\*-eh<sub>2</sub>*) declension. The situation was further complicated by cross-paradigmatic borrowings of endings and, in particular, by the borrowing of endings from pronominal inflection. The genitive plural of *ā*-stem and *o*-stem nouns and adjectives is a case in point. The inherited genitive plural ending *-um* survived in a few *o*-stem nouns, e.g., *deum* ‘gods’, m. gen. pl., but most *o*-stems inflected with *-rum* (with lengthening of the stem vowel), an ending that was modeled on the genitive plural of *ā*-stem nouns, e.g., *casārum* → *deōrum*. But the ending *-rum*, from earlier *\*-som*, was not part of the original *ā*-stem paradigm; it was borrowed from pronominal inflection.

The endings for Classical Latin consonant-stem nouns are given in Table 6.2. The inherited athematic endings, although altered by sound change in many cases, survived in the singular; the Latin ablative was formally the locative ending *\*-i*. In the plural the nominative ending *-ēs* was borrowed from the *i*-stems, where the full-grade stem and the ending fused together to produce a long vowel (*\*-ey-es* > *-ēs*). The Latin accusative plural

ending  $-\bar{e}s$  developed regularly from  $*-\eta s$ . The dative/ablative ending  $-ibus$  had its initial  $-i$  from the  $i$ -stems.

TABLE 6.2 CLASSICAL LATIN CONSONANT-STEM ENDINGS

	Singular		Plural	
	m./f.	n.	m./f.	n.
nominative	$-s$	$-\emptyset$	$-\bar{e}s$	$-a$
vocative	$-s$	$-\emptyset$	$-\bar{e}s$	$-a$
accusative	$-em$	$-\emptyset$	$-\bar{e}s < *-\eta s$	$-a$
genitive	$-is$		$-um$	
dative	$-\bar{i}$		$-ibus$	
ablative	$-e$		$-ibus$	

The endings of the Classical Latin  $o$ -stem declension nouns and adjectives are listed in Table 6.3.

TABLE 6.3 CLASSICAL LATIN O-STEM ENDINGS

	Singular		Plural	
	m.	n.	m.	n.
nominative	$-s$	$-m$	$-\bar{i}$	$-a$
vocative	$-e$	$-m$	$-\bar{i}$	$-a$
accusative	$-m$	$-m$	$-\bar{o}s$	$-a$
genitive	$-\bar{i}$		$-\bar{o}rum$	
dative	$-\bar{o}$		$-\bar{i}s$	
ablative	$-\bar{o}$		$-\bar{i}s$	

Historically, the endings of the nominative, accusative, and dative singular were the same as the athematic endings, although the dative is the result of the contraction of stem vowel and ending, as noted above. The thematic vocative singular was endingless; word-final  $-e$  is the  $e$ -grade of the thematic vowel, e.g., *puere* ‘boy’ (Plautus), cf. Umbrian *tefre* ‘Tefer’. In Very Old Latin and in Faliscan the  $o$ -stems were inflected with two genitive singular endings:  $-sio$  (VOL UALESIOSIO ‘Valerius [personal name]’, m. gen. sg.; Old Faliscan **uo<I>tenosio** ‘Voltenos [personal name]’, m. gen. sg.) and  $-\bar{i}$  (Old Faliscan **titi** ‘Titus’, m. gen. sg.). The functional distinction between the two, if any, is not clear. The ending  $-sio$  does not appear to have survived much beyond the 6th century BCE in Latin, although an Old Latin form attested on an inscription from Ardea may continue  $-sio$  (ΠΙΤΟΙΟ /titojjo/ ‘Titus’, Ardea). The personal name *Mettoeo Fufetioeo* ‘Mettius Fufetius’, which appears in Ennius (Ann. 2, 139 [Warmington]), was probably modeled on the Greek epic genitive  $-\omicron\iota\omicron$ . The ending  $-\bar{i}$  is not a Latino-Faliscan innovation; it is also found in Celtic and Messapic. The PIE thematic declension had a distinct ablative singular ending  $*-\bar{o}d$ . The ending is found in Old Latin, e.g., **GNAIUOD** ‘Gnaeus [personal name]’ and **POPPLICOD** ‘of the people’, and in Oscan, e.g., **sakaraklúd** ‘sanctuary’. Classical Latin  $-\bar{o}$  is the result of the loss of word-final  $-d$  following a long vowel. The expected nominative plural ending is  $-\bar{o}s < *-\bar{o}-es$ . This ending is found in Sabellic languages, e.g., Oscan **núvlanús** ‘inhabitants of Nola’, m. nom. pl., where it was extended to pronominal forms, e.g., Oscan *iusc* ‘they’, m. nom. pl. Latin  $o$ -stems, on the other hand, had the pronominal ending  $*-\bar{o}y$ , which developed regularly to  $-\bar{i}$ , thus *uirí* ‘men’, m. nom. pl. The dative/ablative ending  $-\bar{i}s$  is, from a PIE perspective, the instrumental ending  $*-\bar{o}ys$ .

Neuter 2nd declension nouns had the same ending for the nominative, vocative, and accusative. In the singular, the ending was *-m*; in the plural, the ending was *-a*.

### Declensional classes

In Latin, nouns were organized into five classes called declensions – a practice that is followed in most descriptions of Sabellic nominal forms as well. The five Latin declensions reflect in large part the stem types attested in other ancient Indo-European languages. The *ā*-stems, historically *\*-eh<sub>2</sub>-*stems, are the 1st declension; *o*-stems are the 2nd declension; consonant-stems and *i*-stems are grouped into the 3rd declension; and *u*-stems are the 4th declension. The *ē*-stems, which are an Italic innovation, are 5th. Examples of inherited inflectional types are given in Tables 6.4 and 6.5; *ē*-stems are in Table 6.7.

TABLE 6.4 CLASSICAL LATIN NOUN PARADIGMS, MASCULINE AND FEMININE GENDERS

singular	<i>ā</i> -stems	<i>o</i> -stems	C-stems	<i>i</i> -stems	<i>u</i> -stems
nominative	<i>mēnsa</i> <sup>1</sup>	<i>lupus</i> <sup>3</sup>	<i>rēx</i> <sup>5</sup>	<i>puppis</i> <sup>7</sup>	<i>tribus</i> <sup>8</sup>
vocative	<i>mēnsa</i>	<i>lupe</i>	<i>rēx</i>	<i>puppis</i>	<i>tribus</i>
accusative	<i>mēnsam</i>	<i>lupum</i>	<i>rēgem</i>	<i>puppim</i> <i>-em</i>	<i>tribum</i>
dative	<i>mēnsae</i>	<i>lupō</i>	<i>rēgī</i>	<i>puppī</i>	<i>tribū</i>
ablative	<i>mēnsā</i>	<i>lupō</i>	<i>rēge</i>	<i>puppī, -e</i>	<i>tribū</i>
genitive	<i>mēnsae</i>	<i>lupī</i>	<i>rēgis</i>	<i>puppis</i>	<i>tribūs</i>
locative	<i>Romae</i> <sup>2</sup>	<i>humī</i> <sup>4</sup>	<i>rūrī</i> <sup>6</sup>	—	—
plural	<i>ā</i> -stems	<i>o</i> -stems	C-stems	<i>i</i> -stems	<i>u</i> -stems
nominative	<i>mēnsae</i>	<i>lupī</i>	<i>rēgēs</i>	<i>puppēs</i>	<i>tribūs</i>
vocative	<i>mēnsae</i>	<i>lupī</i>	<i>rēgēs</i>	<i>puppēs</i>	<i>tribūs</i>
accusative	<i>mēnsam</i>	<i>lupōs</i>	<i>rēgēs</i>	<i>puppīs</i>	<i>tribūs</i>
dative	<i>mēnsīs</i>	<i>lupīs</i>	<i>rēgibus</i>	<i>puppibus</i>	<i>tribibus</i>
ablative	<i>mēnsīs</i>	<i>lupīs</i>	<i>rēgibus</i>	<i>puppibus</i>	<i>tribibus</i>
genitive	<i>mēnsārum</i>	<i>lupōrum</i>	<i>rēgum</i>	<i>puppium</i>	<i>tribuum</i>

<sup>1</sup> ‘table’, <sup>2</sup> ‘at Rome’, <sup>3</sup> ‘wolf’, <sup>4</sup> ‘on the ground’, <sup>5</sup> ‘king’, <sup>6</sup> ‘in the country’, <sup>7</sup> ‘stern of a boat’, <sup>8</sup> ‘tribe’.

Declensions 2, 3, and 4 also included nominal forms that were neuter gender. The inflectional patterns differed from the masculine and feminine in the nominative, vocative, and accusative, singular and plural. Partial paradigms are in Table 6.5.

TABLE 6.5 CLASSICAL LATIN NOUN PARADIGMS, NEUTER

singular	<i>o</i> -stems	C-stems	<i>i</i> -stems	<i>u</i> -stems
nom./voc./acc.	<i>iugum</i> <sup>1</sup>	<i>genus</i> <sup>2</sup>	<i>mare</i> <sup>3</sup>	<i>genū</i> <sup>4</sup>
plural				
nom./voc./acc.	<i>iuga</i>	<i>genera</i>	<i>maria</i>	<i>genua</i>

<sup>1</sup> ‘yoke’, <sup>2</sup> ‘kind’, <sup>3</sup> ‘sea’, <sup>4</sup> ‘knee’.

Masculine gender, 2nd declension nouns and adjectives come in a couple of subtypes. The regular type ends in *-us* < *\*-os* in the nominative singular, e.g., *lupus* ‘wolf’, *filius* ‘son’. Nouns and adjectives whose stems ended in *\*-ro-* lost the stem vowel in the



nominative singular by regular sound change, e.g., *uir* ‘man’, *puer* ‘boy’, *sacer* ‘sacred’, etc. Nominative case forms such as *ager* ‘field’ and *sacer* ‘holy’ developed in the following manner: \**agros* > \**agrs* > \**agers* > *ager*. These changes account for the alternations in *-ro*-stem paradigms, e.g., *ager*, m. nom. sg., but *agrī*, gen. sg., etc.; *sacer*, m. nom. sg., but *sacrī*, gen. sg., etc.; cf. VOL SAKROS, m. nom. sg.

The Latin 3rd declension encompasses PIE consonant-stems and *i*-stems. Numerous declensional subclasses can be recognized based on the final sound of the stem; inflectional peculiarities are associated with each class.

1. Stems ending in stop consonants. This class includes a small group of root nouns of the agent or action type, e.g., *rēx*, *rēgis* ‘king’, *pēs*, *pedis* ‘foot’, *dux*, *ducis* ‘leader’, *lūx*, *lūcis* ‘light’. Compounds whose final member was a root noun also belong here, e.g., *auceps*, *aucipis* ‘birdcatcher’ < \**awi*-kaps, as do nouns formed by the derivational suffixes ending in *-tāt* and *-tūt*, e.g., *nouitās*, *nouitātis* ‘newness’, and *uirtūs*, *uirtūtis* ‘manliness’.
2. *s*-stems. The neuter *s*-stems rhotacize medial *-s*, thus *genus* ‘family’, n. nom./acc. sg., but *generis*, gen. sg., *generī*, dat. sg. Some nouns, like *genus*, *generis*, preserve the original *-o/e* ablaut in the suffix (\**o* > *u* in closed final syllables). In other nouns, ablaut was eliminated in favor of the vowel of the nominative/accusative singular, e.g., *corpus* ‘body’, *corporis*. Masculine and feminine *s*-stems were in the process of shifting to the *r*-stem declension in Classical Latin thanks to the analogical introduction of the *-r* of the oblique stem into the nominative singular, e.g., *honōs*, *honōris* → *honor*, *honōris*.
3. *r*-stems. Nouns of familial relationship preserved remnants of ablaut, e.g., *pater*, m. nom. sg., but *patris*, gen. sg. Agent nouns formed by means of the suffix *-tōr* made up the largest constituency in this class, e.g., *uictor*, *uictōris* ‘winner’.
4. *l*-stems. Latin had a few *l*-stem nouns, e.g., *sāl* ‘salt’, *salis*, gen. sg., and *sōl* ‘sun’, *sōlis*, gen. sg. The word for ‘sun’ was in origin a heteroclitic *l/n*-stem. The noun compound *cōnsul* ‘consul’ also belongs to this type. The root is \**sel-* ‘take’.
5. *n*-stems. De-verbal neuter nouns ending in *-men*, *-minis* were very common in Latin, and they were productive in Classical Latin, e.g., *carmen* ‘song’, *carminis*, gen. sg. Animate nouns in *-mō*, *-mōnis* belong to this class as well, e.g., *sermō* ‘speech’, m. nom. sg., *sermōnis*, gen. sg. A small number of nouns preserved traces of ablaut, e.g., *carō* ‘flesh’, f. nom. sg., *carnis*, gen. sg., cf. Umbrian *karu*, f. nom. sg., *karne*, dat. sg. Nouns of the *uirgō*-type also have a distinct stem *uirgin-* outside of the nominative singular, e.g., *uirginis*, f. gen. sg.
6. *m*-stem. Latin has a single *m*-stem, *hiēms* ‘winter’, which, unlike *n*-stems, has a sigmatic nominative singular, /*hiēmps*/. The *p* was epenthetic. The noun belonged to the PIE root class.
7. Heteroclitic *r/n*-stems. Latin preserves a few neuter *r/n*-stems, although the class is moribund and the nouns that do survive show various types of paradigmatic leveling, e.g., *femur* ‘thigh’, *feminis*, gen. sg., but also *femoris*, gen. sg. (form attested in Cicero!); *iecur*, *iocur* ‘liver’, nom. sg., *iecoris*, *iocineris*, *iecinoris*, all gen. sg.
8. *i*-stems. The number of “true” *i*-stems, that is to say, nouns whose paradigms preserved *i*-stem inflection throughout, is relatively small (see Table 6.3). Many *i*-stems have been attracted into consonant-stem inflection to varying degrees.

From an IE perspective, the *i*-stems did not inflect in the manner of consonant-stems. However, sound change and analogy moved *i*-stems, particularly those whose stem vowel was syncopated in the nominative singular, e.g., *mors*, *mortis* ‘death’, *gēns*, *gentis* ‘clan’, etc., in the direction of consonant-stem inflection, which in turn moved consonant-stem nouns in the direction of *i*-stems, e.g., *cīuitātium* ‘states’, f. gen. pl., *cīuitātīs* ‘states’,

acc. pl. The result is a class of 3rd declension nouns in Classical Latin whose inflection lies between that of “pure” consonant-stems and “pure” *i*-stems, and which is sometimes referred to, though somewhat infelicitously, as “mixed-stem” inflection. Nouns of this class typically had variant forms of the accusative singular, *-em* and *-im*; ablative singular, *-e* and *-ī*; accusative plural, *-ēs* and *-īs*; and, less frequently, the genitive plural, *-um* and *-ium*. Examples of this inflection are cited in Table 6.6.

TABLE 6.6 CLASSICAL LATIN ‘MIXED’ DECLENSION

singular				
nominative	<i>gēns</i> ‘family’	<i>fōns</i> ‘spring’	<i>mors</i> ‘death’	<i>pars</i> ‘portion’
vocative	<i>gēns</i>	<i>fōns</i>	<i>mors</i>	<i>pars</i>
accusative	<i>gentem</i>	<i>fontem</i>	<i>mortem</i>	<i>partem, -im</i>
dative	<i>gentī</i>	<i>fontī</i>	<i>mortī</i>	<i>partī</i>
ablative	<i>gente</i>	<i>fonte, -ī</i>	<i>morte</i>	<i>parte, -ī</i>
genitive	<i>gentis</i>	<i>fontis</i>	<i>mortis</i>	<i>partis</i>
plural				
nominative	<i>gentēs</i>	<i>fontēs</i>	<i>mortēs</i>	<i>partēs</i>
vocative	<i>gentēs</i>	<i>fontēs</i>	<i>mortēs</i>	<i>partēs</i>
accusative	<i>gentīs, -ēs</i>	<i>fontīs</i>	<i>mortēs</i>	<i>partēs, -īs</i>
dative	<i>gentibus</i>	<i>fontibus</i>	<i>mortibus</i>	<i>partibus</i>
ablative	<i>gentibus</i>	<i>fontibus</i>	<i>mortibus</i>	<i>partibus</i>
genitive	<i>gentium</i>	<i>fontium, -um</i>	<i>mortium</i>	<i>partium, -um</i>

Oscan maintained the distinction between *i*-stems and consonant-stems in most case forms, including the nominative plural, e.g., **trīs** ‘three’ < \**treyes*, but **meddiss** ‘magistrate’, m. nom. pl. < \**meddikēs*, with syncope of short \*-e in the final syllable. Even so, these stem classes do exhibit cross-paradigmatic borrowings. For example, the *i*-stem genitive singular ending *-eis* was adopted in consonant-stem inflection in all Sabellic languages.

The 5th declension is an Italic development, though few forms, apart from Umbrian **ri** ‘matter’, dat. sg., are attested for this class in Sabellic. In Latin, the 5th declension nouns *diēs* ‘day’, *rēs* ‘property’, *spēs* ‘hope’, and *fidēs* ‘trust’ come historically from different PIE nominal classes, but sound change and analogical change have led to fully elaborated paradigms. Examples in are Table 6.7.

TABLE 6.7 CLASSICAL LATIN *Ē*-STEMS

singular		<i>ē</i> -stems	
nominative		<i>diēs</i> ‘day’	<i>rēs</i> ‘property’
vocative		<i>diēs</i>	<i>rēs</i>
accusative		<i>diem</i>	<i>rem</i>
dative		<i>diei</i>	<i>rēi, rei</i>
ablative		<i>diē</i>	<i>rē</i>
genitive		<i>diēi, diē, diī</i>	<i>rēi, rei</i>
plural			
nominative		<i>diēs</i>	<i>rēs</i>
vocative		<i>diēs</i>	<i>rēs</i>
accusative		<i>diēs, diīs</i>	<i>rēs</i>
dative		<i>diēbus</i>	<i>rēbus</i>
ablative		<i>diēbus</i>	<i>rēbus</i>
genitive		<i>diērum</i>	<i>rērum</i>

## Diachronic developments

The Italic languages shared one major innovation in the nominal system, and that was the development of ablative singular forms ending in a long vowel + *-d*. The ending was modeled on the *o*-stem ablative *-ōd*, and it was imported into the *ā*-stems, *i*-stems, and *u*-stems, which yielded ablatives distinct from their genitive forms. The 5th declension nouns adopted this ending as well. A few consonant-stem ablatives were generated following this model, e.g., OL [C]ONSOLEO ‘consul’, m. abl. sg., LEGED ‘law’, f. abl. sg.; in other cases consonant-stems opted for the ablative of the *i*-stem inflection, e.g., OL OPID ‘resources’, f. abl. sg., BOVID ‘cow’, m./f. abl. sg.

## Adjectives

The most common type of adjective declension paired masculine and neuter forms of the 2nd declension with feminine forms of the 1st, e.g., *bonus*, *-a*, *-um* ‘good’. In consonant-stem declension, distinctions in gender were not formally expressed in the nominative singular, e.g., *ferēns* ‘carrying’, *ferōx* ‘fierce’, *duplex* ‘double’, *ingēns* ‘huge’ (*ingentis*, gen. sg.), *uetus* ‘old’ (*ueteris*, gen. sg.), etc. For most *i*-stems, masculine and feminine nominatives were distinct from the neuter, e.g., *facilis* ‘easy’, m./f., *facile* n. In paradigms such as *ācer*, *ācris*, *ācre* ‘sharp’, a three-way distinction in gender was created in the nominative singular. \**ākris* developed regularly to *ācer* in Old Latin; *ācris* was reintroduced, based perhaps on the model of *facilis*, etc. Interestingly enough, in Old Latin *ācris* is found as a masculine form and *ācer* as a feminine form. The match-up of gender and form found in Classical Latin is the result of standardization on the part of elite Latin writers.

Latin did not have *u*-stem adjectives; they were transferred to the *i*-stem inflection, e.g., *suāuis*, *suāue* ‘sweet’ ← PIE \**swādu-*, cf. Greek ἡδύς. Present active participles in *-nt-*, originally consonant-stems in their inflectional pattern, were attracted into the orbit of the *i*-stem inflectional type. The ablative singular ends in *-ī* unless the participle is used as a substantive, e.g., *portantī* ‘carrying’, m. dat. sg. The genitive plural is *-ium*, e.g., *portantium* ‘carrying’, m. gen. pl.

Gradable adjectives formed the comparative by means of the suffixes *-iōr* < \**-yōs* (m./f.) and *-ius* < \**-yos* (n.), e.g., *melior*, *meliōris*, ‘better’, m./f., *melius*, *meliōris*, n. As was the case for *s*-stem masculine and feminine nouns, so in the comparative the oblique stem suffix, whose medial *-r* was the result of rhotacism, was introduced into the nominative singular. The superlative suffix was *-issimus*, *-a*, *-um* with 1st and 2nd declension inflection. The superlative of a few *i*-stems, whose stem vowel was syncopated, ended in *-limus* due to assimilation of *-ls* to *-ll*, e.g., \**fakilisomos* > \**fakilsomos* > *facillimus* ‘most easy’.

## Pronouns

Pronominal declension includes personal pronouns, which were not inflected for gender; anaphoric/demonstrative pronouns, which were inflected for gender; and interrogative and relative pronouns, which were also inflected for gender.

The personal pronouns were idiosyncratic in their inflection, and this is reflected in the paradigms in Table 6.8. Case forms of the 1st and 2nd person personal pronouns were built on multiple stems, e.g., *egō*, *me-/mē*, *nos-/nōs*. Some case endings, for example, the dative singular endings *-hi* and *-bi*, and the dative/ablative plural *-bīs*, were unique to

these paradigms. The genitive singular and plural forms were drawn from the possessive adjective paradigms. The Old Latin genitives *mīs* ‘of me’ and *tīs* ‘of you’ are the inherited enclitic forms \**mey* and \**tey*, extended by the *-s* of the genitive. The reflexive pronoun had the same forms in the singular and plural.

TABLE 6.8 LATIN PERSONAL PRONOUNS

	1st person	2nd person	Reflexive
singular			
nominative	<i>ego</i> (OL <i>egō</i> )	<i>tū</i>	—
accusative	<i>mē</i> (OL <i>mēd</i> )	<i>tē</i> (OL <i>tēd</i> )	<i>sē</i> (OL <i>sēd</i> )
dative	<i>mihi</i> (OL <i>mihī</i> )	<i>tibi</i> (OL <i>tibī</i> )	<i>sibi</i> (OL <i>sibī</i> )
ablative	<i>mē</i> (OL <i>mēd</i> )	<i>tē</i> (OL <i>tēd</i> )	<i>sē</i> (OL <i>sēd</i> )
genitive	<i>meī</i> ( <i>mīs</i> )	<i>tūī</i> ( <i>tīs</i> )	<i>suī</i>
plural			
nominative	<i>nōs</i>	<i>uōs</i>	
accusative	<i>nōs</i>	<i>uōs</i>	
dative	<i>nōbīs</i>	<i>uōbīs</i>	
ablative	<i>nōbīs</i>	<i>uōbīs</i>	
genitive	<i>nostrum, nostrī</i>	<i>uestrum, uestrī</i>	

The personal pronouns attested in Sabellic languages match up tolerably well with those in Latin, e.g., South Picene **ekū** /egō/, nom. sg., Umbrian **mehe**, dat. sg., Umbrian **miom** /mēmō/, acc. sg.; Oscan **tiium**, nom. sg., **tif[ei]**, **t(i)feī**, dat. sg.; South Picene **tefeī**, dat. sg.; Umbrian **tefe**, dat. sg.; Paelignian **uus**, nom. pl.; Oscan **sifeī**; Paelignian **sefei**, dat. sg. Paelignian **uus** ‘to you’, dat. pl., points to a prehistoric \**uōfos*. If this was the Proto-Italic form of the dative/ablative, then the endings of Latin *nōbīs* and *uōbīs* were reformed to \**nōbeis* and \**uōbeis* based the dative singular. In Very Old Latin and Old Faliscan, accusative and ablative singulars ended in *-d*, e.g., VOL MED ‘me’, acc. sg., Old Faliscan **med**, acc. sg. Why the accusative forms ended up with word-final *-d* is not easily explained. In contrast to Latin and Faliscan, the particle *-om* was added to the accusative singular forms in Sabellic, e.g., Umbrian **miom** ‘me’, **tiom** ‘you’. The 1st person accusative singular was remade in Venetic by contamination with the nominative, \**egō* : \**mē* → **ego** : **meḡo**.

Gender-bearing pronouns represent a variety of distinct stem types. The stems \**e*, \**ey*, and \**i* united to form the paradigm of anaphoric pronouns *is*, *ea*, *id* ‘he, she, it’, in Latin, and *izic* ‘he’, **iúk** ‘she’, **ídik** ‘it’ in Oscan. The Latin demonstrative *hic*, *haec*, *hoc* ‘this’ (close to speaker) is not attested as such in any other language, though it is often thought to be from the same etymological source as the Sanskrit particle *gha* ‘certainly’. *iste*, *ista*, *istud* ‘this’ (close to addressee) was formed from a particle \**es-* + demonstrative \**to-/tā-*. Umbrian **este**, ‘this’, n. acc. sg., South Picene **estas**, f. nom. pl. (?), and Presamnite **estam**, f. acc. sg., are cognate. The *i*-vocalism in Latin is due to the influence of *is*. *ille*, *illa*, *illud* ‘that’ (distant from speaker) is the Classical Latin form that replaced OLat. *olle/ollus*, cf. Oscan **úlleis**, ‘that’, m. gen. sg. The *i*-vocalism here may be attributed once again to *is* and perhaps to *iste* as well. These paradigms shared an inflectional peculiarity by which the genitive and dative singular forms were the same regardless of gender, e.g., *huius*, *huic*; *illius*, *illī*; *istius*, *istī*.

The deictic particle *-c(e)* was often added to inflected pronominal forms in Latin and Sabellic. In Latin, demonstrative pronouns to which the particle had been added were incorporated into the paradigms in some case forms, e.g., *hic*, *haec*, *hoc* ‘this, these’,

nom. sg.; *hunc, hanc, hoc*, acc. sg. The same was true also for Oscan and Umbrian, e.g., Oscan **íúk** ‘she’ < \*eyā-ke.

The paradigm of the Latin relative pronoun is cited in Table 6.9. Some forms (*quem*, m. sg. acc.) go back to the PIE interrogative/indefinite stem \*k<sup>w</sup>i-, others (*quod*, n. sg. nom./acc.) to the interrogative stem \*k<sup>w</sup>o-. The PIE relative pronoun \*yo- was lost in Italic, and the relative paradigm was composed of the stem \*k<sup>w</sup>o- and the stem \*k<sup>v</sup>i-. The use of the interrogative/indefinite pronouns with relative function is to be attributed to Proto-Italic; the paradigms in Oscan and Umbrian are similarly formed. In Classical Latin the relative pronoun was distinct from the interrogative and indefinite only in the nominative singular. The masculine and feminine relative pronouns were augmented by the deictic particle -ī, e.g., Latin *quī* ‘who’, m. nom. sg., < \*k<sup>w</sup>o-ī, cf. Oscan **puī**.

TABLE 6.9 CLASSICAL LATIN RELATIVE PRONOUN

	relative		
	masculine	feminine	neuter
nominative	<i>quī</i>	<i>quae</i>	<i>quod</i>
accusative	<i>quem</i>	<i>quam</i>	<i>quod</i>
dative	<i>cuī</i>	<i>cuī</i>	<i>cuī</i>
ablative	<i>quō</i>	<i>quā</i>	<i>quō</i>
genitive	<i>cuius</i>	<i>cuius</i>	<i>cuius</i>
plural			
nominative	<i>quī</i>	<i>quae</i>	<i>quae</i>
accusative	<i>quōs</i>	<i>quās</i>	<i>quae</i>
dative	<i>quibus</i>	<i>quibus</i>	<i>quibus</i>
ablative	<i>quibus (quīs)</i>	<i>quibus</i>	<i>quibus</i>
genitive	<i>quorum</i>	<i>quārum</i>	<i>quōrum</i>

## VERB MORPHOLOGY

Verbs were inflected for tense (pres., impf., fut., perf., pluperf., fut. perf.), mood (ind., subj., imp.), voice (act., pass.), person (1st, 2nd, 3rd), and number (sg., pl.).

The Italic languages, excluding Venetic, for which there is too little evidence, reorganized the verb system inherited from PIE. In Italic, the primary division was into two basic stems, generally referred to as the imperfective and the perfective. Three tense formations were constructed to each stem: to the imperfective, there was a present, an imperfect, and a future; to the perfective, a perfect, a pluperfect, and a future perfect. This organizational schema is set out for Latin and Oscan in Table 6.10.

TABLE 6.10 ORGANIZATION OF VERB SYSTEM

Latin	<i>imperfective</i>	<i>perfective</i>	
present	<i>damus</i> ‘we give’	perfect	<i>dedimus</i> ‘we gave’
imperfect	<i>dabāmus</i> ‘we were giving’	pluperfect	<i>dederāmus</i> ‘we had given’
future	<i>dabimus</i> ‘we will give’	future perfect	<i>dederimus</i> ‘we will have given’
Oscan	<i>imperfective</i>	<i>perfective</i>	
present	<i>didet</i> ‘he gives’ (Vestinian)	perfect	<i>deded</i> ‘he gave’
imperfect	<i>fufans</i> ‘they were’	pluperfect	(unattested)
future	<i>didest</i> ‘he will give’	future perfect	<i>tribarakattust</i> ‘he will have built’

In the imperfective, verbs were organized into four present-tense paradigmatic classes, commonly referred to as conjugations: conjugation 1, *ā*-stems; conjugation 2, *ē*-stems; conjugation 3, *i*-stems; and conjugation 4, *ī*-stems. Conjugation 3 had two subtypes: a regular 3rd conjugation, in which the stem vowel *-i* did not appear before endings of the 1st singular and 3rd plural (see below, *regō* ‘guide’); and the 3*īō* conjugation, in which the stem vowel *-i* appeared in all forms of the present (see below, *faciō* ‘make’). The Latin conjugational classes, present tense, are displayed in Table 6.11. Verbs like *ferō* ‘I carry’, *uolō* ‘I wish’, *sum* ‘I am’, *eō* ‘I go’, etc., which do not fit neatly into one of the five patterns, are described as “irregular”.

TABLE 6.11 LATIN CONJUGATION CLASSES

	<i>ā</i> -stems	<i>ē</i> -stems	<i>i</i> -stems	<i>i</i> -stems ( <i>-iō</i> )	<i>ī</i> -stems
1.	<i>portō</i>	<i>moneō</i>	<i>regō</i>	<i>faciō</i>	<i>sentīō</i>
2.	<i>portās</i>	<i>monēs</i>	<i>regis</i>	<i>facis</i>	<i>sentīs</i>
3.	<i>portat</i>	<i>monet</i>	<i>regit</i>	<i>facit</i>	<i>sentit</i>
1.	<i>portāmus</i>	<i>monēmus</i>	<i>regimus</i>	<i>facimus</i>	<i>sentīmus</i>
2.	<i>portātis</i>	<i>monētis</i>	<i>regitis</i>	<i>facitis</i>	<i>sentītis</i>
3.	<i>portant</i>	<i>monent</i>	<i>regunt</i>	<i>faciunt</i>	<i>sentiunt</i>

The imperfect and future tenses were formed on the imperfective stem. The imperfect tense suffix was *-bā* for all conjugations, e.g., 1st conjugation *portābās* ‘you were carrying’, 2nd conjugation *monēbās* ‘you were warning’. Third and 4th conjugation verbs add the imperfect tense suffix to a stem augmented by *-ē*, e.g., 3rd conjugation *regēbās* ‘you were guiding’, 3*īō* conjugation *faciēbās* ‘you were making’, 4th conjugation *sentīēbās* ‘you were feeling’. The future tense had two formants: (1) *-b/bi-*, which was added to the present stems of conjugations 1 and 2, e.g., *portābō* ‘I will carry’, *monēbis* ‘you will warn’; and (2) *-ā* (1st sg. only)/*-ē*, which was added to the present stems of conjugations 3 and 4, e.g., *dūcam* ‘I will lead’, *faciēs* ‘you will make’, *sentīēmus* ‘we will feel’. In Old Latin, imperfects of 4th conjugation verbs could be formed from the stem un-augmented by *-ē*, e.g., *exaudībat* ‘he was listening’, *seruībās* ‘you were preserving’. *-b/bi-*formations could also be formed to verbs of the 4th conjugation, e.g., *dormībō* ‘I will sleep’. These formations are rare in Classical Latin literature. The future attested in Faliscan corresponds to the *-b/bi-*formation in Latin, e.g., 2nd conjugation **carefo** ‘I will be lacking’.

The stem of the perfective active was formed in a number of ways from a synchronic point of view: (1) by suffixation of *-u* /*w*/, e.g., *portāuit* ‘he carried’, *monuit* /*monuwit*/ ‘he warned’ < \**monewed*, *sentīuit* ‘he felt’; (2) by reduplication, *dedit* ‘he gave’, beside present tense *dat*, cf. Faliscan **pe:parai** ‘I produced’, **fifiked** ‘he fashioned’, **fi[fi]qod** ‘they fashioned’; (3) by lengthening of the stem vowel, *uēnit* ‘he came’, *rūpit* ‘he burst’, beside present tense *uenit*, *rupit*; (4) by altering the quality of the stem vowel, accompanied by lengthening, e.g., *fēcit* ‘he made’, to *facit* ‘he makes’; (5) by suffixation of *-s*, e.g., *dīxit* /*dī:ksit*/ ‘he said’, beside *dīcit* ‘he says’. In some cases, the imperfective and perfective stems were distinguished only by the personal endings, e.g., present tense *uertō* ‘I turn’, but perfect tense *uertī* ‘I turned’.

Perfective active stems with the suffix *-u* were generally built from the imperfective stem of secondary verbs, most of which were denominatives ending in a long vowel, e.g., *cūrā-* ‘to take care to’ → *cūrā-u-*. Other perfective formations were built to verbal roots

or to verbal stems distinct from the stem of the imperfective, so that the morphological relationship between the imperfective stem and the perfective stem was not always predictable, particularly outside of the 1st and 4th conjugations, e.g., *monē-* ‘warn’, but *monu-* /*monuw-*/ built from the stem \**mone-*.

Alongside the regular *ā*-conjugation perfects, e.g., *portāuisti* ‘you carried’, there were also “syncopated” or “contracted” perfect forms, e.g., *portāstī*. Such forms were common in Republican Latin literature, particularly in poetry, where they provided metrically convenient alternatives to longer forms. But despite the labels, these forms were not the result of syncope or contraction. Rather, they originated as sigmatic (aorist) forms to denominatives, and they were eventually incorporated into the paradigms of the *u*-perfects, e.g., *portāstī* = *portā-s-stī*; *portārun*t < \**portā-s-ont*.

The pluperfect suffix for active voice was *-erā-*; the future perfect suffix was *-er/-eri-*. Both suffixes were added to the perfective stem, e.g., *portāu-erā-s* ‘you had carried’, *dix-erā-s* ‘you had said’; *portāu-er-ō* ‘I will have carried’, *monu-eri-s* ‘you will have warned’.

## Mood

In addition to the indicative mood, Italic languages had subjunctive mood forms corresponding to all tenses in the system with the exception of the future and future perfect. Forms for each conjugational class are listed in Table 6.12.

TABLE 6.12 SUBJUNCTIVE MOOD (2 SG. ACT.)

	<i>ā</i> -stems	<i>ē</i> -stems	<i>i</i> -stems	<i>i</i> -stems(- <i>iō</i> )	<i>ī</i> -stems
present	<i>portēs</i>	<i>moneās</i>	<i>dīcās</i>	<i>faciās</i>	<i>sentīās</i>
imperfect	<i>portārēs</i>	<i>monērēs</i>	<i>dīcerēs</i>	<i>facerēs</i>	<i>sentīrēs</i>
perfect	<i>portauerīs</i>	<i>monuerīs</i>	<i>dīxerīs</i>	<i>fēcerīs</i>	<i>sentīuerīs</i>
pluperfect	<i>portauissēs</i>	<i>monuissēs</i>	<i>dixissēs</i>	<i>fēcissēs</i>	<i>sentīuissēs</i>

The suffixes marking subjunctive mood are the following: present tense *-ā* for conjugations 2, 3, and 4, and present tense *-ē* for conjugation 1; imperfect tense *-rē*; perfect tense *-erī-*; and pluperfect tense *-issē-*. As the forms in Table 6.11 shows, for conjugation 1 the subjunctive suffix *-ē* replaces the stem vowel *-ā* characteristic of the present indicative. The final vowel of the perfect tense suffix, to judge from the variation found in Classical Latin poetry, could be either long or short.

The imperative, which was formed on the imperfective stem, was signaled by distinct endings rather than by a special mood suffix. Latin had two sets of imperative endings, one for the “regular” imperative and one for the so-called future imperative. Active voice formations are 2 sg. *age* vs. 2/3 sg. *agitō*; 2 pl. *agite* vs. 2/3 pl. *agitōte*. Functionally, the so-called future imperative was restricted primarily to legal and juridical contexts.

## Voice and personal endings

In Latin, verbs were inflected for active and passive voice. Middle voice as a grammatical category had for the most part disappeared, although a small number of verbs permitted middle voice readings, e.g., *lauātur* ‘washes herself, bathes’. The so-called deponents continued PIE *media tantum* verbs, e.g., *sequitur* ‘he follows’, cf. Greek *ἕπεται*, Sanskrit *sacate*.

There were three sets of personal endings for verbs (see Table 6.13). One set was for active voice, and another set was for passive voice and for deponents. The active voice

had two endings for the 1st singular active: *-ō* for present, future, and future perfect, and *-m* for imperfect, pluperfect, and all tenses of the subjunctive. First singular *-m* in the present tense of the verb ‘be’, *sum*, ESOM ‘I am’ is a remnant of the primary athematic ending *\*-mi*, which has otherwise been lost in Italic. The perfect tense had its own special set of forms in the active.

TABLE 6.13 PERSONAL ENDINGS

	active	passive/deponent	perfect active
1 sg.	<i>-ō, -m</i>	<i>-or, -r</i>	<i>-ī</i>
2 sg.	<i>-s</i>	<i>-re, -ris</i>	<i>-istī</i>
3 sg.	<i>-t</i>	<i>-tur</i>	<i>-it</i>
1 pl.	<i>-mus</i>	<i>-mur</i>	<i>-imus</i>
2 pl.	<i>-tis</i>	<i>-mini</i>	<i>-istis</i>
3 pl.	<i>-nt/-unt</i>	<i>-ntur, -untur</i>	<i>-erunt, -erunt, -ēre</i>

The perfect, pluperfect, and future perfect passives were periphrastic formations made up of the perfect passive participle and a finite form of the verb ‘be’. The participle was inflected for gender, number, and case; the inflectional features were determined by agreement with the subject of the verb. Person and number of the subject were also marked on the verb ‘be’. Examples are cited in Table 6.14.

TABLE 6.14 PERFECT PASSIVES

perfect	<i>portāta es</i> ‘you (f.) were carried’
pluperfect	<i>portātū erāmus</i> ‘we (m.) had been carried’
future perfect	<i>portātus erō</i> ‘I (m.) will have been carried’

The paradigms listed in Table 6.11 are evidence for a regular alternation in the length of the stem-final vowel typical of all Latin verbal paradigms. The morphophonemic alternations were the result of sound changes that took place in Old Latin. The first change shortened long vowels when they stood before word-final obstruents, except for *-s*, e.g., *portāt* > *portat*, *faciām* > *faciam*, but *portās*, *portāmus*, etc. The second change shortened long vowels when they stood before another vowel (so-called *vocalis ante vocalem corripitur*), e.g., *\*monēō* > *moneō*, *\*monēās* > *moneās*, *\*sentīēs* > *sentiēs*. From a synchronic point of the view, the stem-final vowels in conjugations 1, 2, and 4 are to be treated as long in underlying forms. In conjugations 3 and 3iō the morphophonemic alternation between stem-final *-i* and stem-final *-e* is best viewed synchronically as a change of *-i* to *-e* before *-r*. This morphophonemic change explains the short *-e* rather than the short *-i* in forms such as *caperēs*, 2 sg. impf. subj. In the underlying representation this verb was /*kapirēs*/. From a diachronic perspective, *caperēs* goes back to *\*kapesēs*.

### Diachronic developments

The present-tense conjugation classes are the result of the convergence of different PIE formations resulting from sound changes and analogical restructurings.

The 3rd conjugation verbs continue PIE or post-PIE thematic formations of various sorts, e.g., Latin *agīt* ‘leads’ < *\*aǵeti*, *dīcit* ‘says’ < *\*deyketi*, *bibit* ‘drinks’ < *\*pibeti*, *poscit* ‘demands’ < *\*pṛksketi*. In Old Latin, the thematic vowels *\*e* and *\*o* changed to *i*



and *u*, e.g., \**dewketi* > \**dowket* > *dūcit*, \**dewkonti* > \**dowkont* > *dūcunt*. Athematic formations, such as nasal infix verbs (e.g., *iungit* ‘joins’), reduplicated presents (e.g., *gignit* ‘gives birth to’), and a few presents to roots ending in a laryngeal (e.g., *sonit* ‘sounds’ < \**swenati* < \**swenh<sub>2</sub>ti*) were added to the thematic conjugation too, though by diverse prehistoric routes.

The 3rd conjugation verbs of the *capiō*-type and 4th conjugation verbs of the *seniō*-type were \*-*ye/yo-* formations built directly to verb roots. The assignment of a verb to either the 3rd or 4th conjugation was determined by the prosodic structure of the root. Prosodically heavy roots and roots ending in a sonorant developed an epenthetic *-i* between the final consonant of the root and the \*-*ye/yo-* suffix. Sound changes then split the verbs into distinct paradigmatic classes, e.g., conjugation 3*iō*, \**kapiēsi* > \**kapes* > *capis* ‘you seize’, but 4th conjugation, \**wenyēsi* > \**wenyēsi* > *uenīs* ‘you come’.

A majority of 2nd conjugation verbs were in origin either iterative/causative formations with *o*-grade of the root and a suffix \*-*eye/o-*, or stative formations in \*-*eh<sub>2</sub>*, e.g., *nocēs* ‘you cause harm to’ < \**nokeyēsi*, *sedēs* ‘you are sitting’ < \**sedeh<sub>1</sub>si*.

The 1st conjugation verbs included denominatives to *ā*-stems, e.g., *curā*- ‘make sure of’ < \**koysā-ye/o-*; factitives to thematic stems, e.g., *nouā*- ‘makes new’ < \**neweh<sub>2</sub>-ye/o-*; and \*-*ye/yo-* formations to roots ending in a laryngeal, e.g., \**swenh<sub>2</sub>-ye/o-* > \**swe-na-ye/o-* > *sonā*- ‘sounds’.

The Italic perfect tense indicatives are the result of the merger, both formally and functionally, of the PIE aorist and the PIE perfect. For most verbs, either an aorist or a perfect was selected as the representative of the Italic perfect, although in earlier periods of Latin both aorist and perfect formations are attested for a few verbs. So, for example, in Old Latin the verb *parcō* ‘I refrain from’ had multiple perfect tense forms: *pepercī*, *parcuī*, *parsī* ‘I refrained from’. The verb *faciō* made a reduplicated perfect and a root aorist in Very Old Latin: *vhe:vhakēd* ‘he made’ and *fecēd*. Latin, Oscan, and Umbrian sometimes selected the same perfective form, e.g., Latin *dedit* ‘gave’, beside Umbrian *dede*. But more often the perfective formation in Latin does not match that of Oscan or Umbrian, e.g., Latin *dixit* ‘said’, sigmatic aorist, beside Umbrian *dirsicust*, reduplicated perfect; and Latin *pepulit* ‘he struck’, reduplicated perfect, beside Umbrian *a(m)pelust* ‘will have struck’, root aorist.

Although the overall structure of the tense system is the same in Italic, the exponents of some “tense” suffixes differed between Latino-Faliscan and Sabellic. For example, the Latin and Faliscan future tense for conjugations 1 and 2 was based on the suffix \*-*b<sup>h</sup>e/o-*, whereas Sabellic continued an *s*-future, in all likelihood, the PIE desiderative, e.g., Oscan *deiuast* ‘he will swear’. The Latin future perfect was a thematic formation added to the perfective formant *-is*, *-eri-* < \*-*is-e/o-*, whereas the Sabellic formation is athematic *-us-*, e.g., Oscan *tribarakattust* ‘he will have built’. The Latin pluperfect suffix was *-er-ā-* < \*-*is-ā-*, which was built with the same *ā*-suffix found in the imperfect of the verb ‘be’, *erās* ‘you were’ < \**es-ā-s*, and in the imperfect suffix *-bā-*. The pluperfect is not attested in Sabellic.

The Latin imperfect formations, both indicative \*-*b<sup>h</sup>ā-* and subjunctive \*-*sē-*, have exact correspondences in the Sabellic languages, e.g., Oscan *fufans* ‘they were’, 3 pl. impf. act., Oscan *patensins* ‘should open’, 3 pl. impf. subj. The present subjunctives of the 1st conjugation have their *ē*-vowel from the full-grade athematic optative suffix, e.g., \**noweh<sub>2</sub>-yeh<sub>1</sub>-* > \**nouāē-* > *nouē-*, a stem that was generalized to plural forms in the paradigm. This subjunctive formant then spread from 1st conjugation athematics to the large class of denominatives in this conjugation, e.g., *cūrē-*, cf. Oscan *deiuaid* ‘he shall swear’, 3 pl. pres. subj. < \**deywā-ē-*. The origin of the *ā*-subjunctive of conjugations 2,

3, and 4 remains a mystery, but the formation is Italo-Celtic in origin and is found in all branches of Italic, e.g., Faliscan **douiad** ‘let him give’, 3 sg. pres. subj.; Oscan **pútiad** ‘be able’, 3 sg. pres. subj.; Oscan *deicans* ‘let them say’, 3 pl. pres. subj.; Oscan **fakiad** ‘let them make’, 3 sg. pres. subj.; and Umbrian *tursiandu* ‘let them be frightened away’, 3 pl. pres. subj.

The distinction between PIE primary and secondary personal endings was, for the most part, eliminated in Latin thanks to the loss of *-i* in the primary endings. In Latin, a distinction remained in the 1st person singular active, where, as noted above, the ending *-m* appeared in past tense formations and the subjunctive, and the ending *-ō* appeared in the present, future, and future perfect. The 3rd person singular active secondary ending *-d* < \**-t* survived in Very Old Latin and Old Latin, e.g., VOL FHE:VHAKED, OL FECID, but was replaced by primary *-t* < \**-ti* before the beginning of the Classical period. The perfect active endings continued the PIE endings for the most part, though they were extended by addition of \**-i*, the hic et nunc particle, in the singular (*-ī* < \**-ay*, *-stī* < \**-stay*, *-it* < *-ī(t)* < \**-ey(t)*) and 3rd person plural (*-ēre* < \**-ēri*). Third singular *-ei* was re-characterized by the addition of *-t*, which developed to *-tī* in Old Latin (OL FUEIT) and then to *-it* in Classical Latin. In addition to *-ēre*, two additional 3rd plural endings are attested: (1) *-ērunt*, which is a blending of *-ēre* and thematic aorist *-unt* ← \**-ond*; and (2) *-erunt*, which is from earlier \**-is-ont*, the thematic aorist ending added to the perfect formant *-is-*. The endings of the Faliscan perfect paradigm appear to be a fusion of thematic aorist and perfect endings, e.g., 1 sg. perfect **pe:parai** ‘I have produced’; 3 sg. aorist **fifiked** ‘he has fashioned’; 3 pl. aorist **f[if]iqod** ‘they have fashioned’. In the Sabellic languages, the primary and secondary endings of the active did not merge, apart from the 2nd person, and perhaps the 1st person plural. Thus, in Oscan and Umbrian the endings are 1 sg. *-u* (Umbrian *stahu* ‘I stand’, 1 sg. pres. act.) vs. **-um** (Oscan **manafum** ‘I entrusted’, 1 sg. perf. act.); **-t** (Oscan **stait** ‘it stands’, 3 sg. pres. act.) vs. **-d** (Oscan **deded**, 3 pl. perf. act.); **-nt** (Oscan **sent** 3 pl. pres. act.) vs. **-ens** (Oscan **dedens** 3 pl. perf. act.). In the 3rd plural, South Picene and Presamnite continued the thematic aorist ending \**-ond* < \**-ont*. These languages did not share the change of \**-nd* to *-ns*, which is found in Oscan and Umbrian, e.g., South Picene **adstúh** ‘they set up’, 3 pl. perf. < \**-ond*, Presamnite *fuffoð* ‘they were’, 3 pl. perf. < \**-ond*.

The endings of the passive voice were built on the active by addition of *-(o)r* or, in the case of the 1st plural, the replacement of *-s* by *-r*, e.g., *-mur* < \**-mor*. The 2nd singular ending *-re* goes back to the PIE secondary middle ending \**-so*. The alternative 2nd singular ending *-ris* is a later formation; final *-s* is the active singular ending. The 2nd plural *-minī* is sometimes taken to be the nominative masculine plural participle; another explanation points to PIE \**-d<sup>h</sup>we*, which received a nasal extension in the manner of Sanskrit 2nd plural active *-thana*, that is to say, \**-d<sup>h</sup>we-ney* > \**-beney* > \**-bney* > \**-mney* > *-minī*. Neither possibility is convincing.

The middle endings in the Sabellic languages did not correspond in all particulars to those in Latin. In Oscan the 3rd singular and 3rd plural suffixes were *-ter* and *-nter*. In Umbrian the *-ter* and *-nter* are used in primary inflection; *-tur* was used in secondary inflection.

### Non-finite formations

The verb system also traditionally includes, in addition to the forms inflected for person and number, a number of non-finite forms. These include participles (*portāns*, *portantis* ‘carrying’), infinitives (*portāre* ‘to carry’, *portāuisse/portāsse* ‘to have carried’), supines

(*perditum* ‘to lose’), gerundives (*ūitandae* ‘worthy of living’), and gerunds (*faciendō* ‘act of making’).

The present active participle was formed by addition of the suffix *-nt-* to the imperfective stem for conjugations 1 and 2, e.g., *portāns* < \*portants, *portantis*, and by addition of *-ent-* for conjugations 3 and 4, e.g., *faciēns*, *facientis*. Gender distinctions were in large part eliminated. The masculine and feminine forms were the same throughout, and the neuter was distinct only in the accusative singular and the nominative/accusative plural, e.g., *facientem*, m./f., vs. *faciēns*, n.; *facientēs*, m./f., vs. *facientia*, n. The neuter form *facientia* shows another innovation of the present participle declension in Latin: *i*-stem inflection in the ablative singular and genitive plural.

Present middle participles in *-minus* < \*-mh<sub>1</sub>no- are preserved in a few words, e.g., *fēmīna* ‘woman’ < \*d<sup>h</sup>eh<sub>1</sub>(i)- ‘give suck’, *alumnus* ‘nursling’ < \*al- ‘nourish’, but these stand outside of the verb system.

The PIE verbal adjective suffix *\*-to-* is the source of the Latin perfect participle, e.g., *dictus* ‘said’, *ductus* ‘led’. Roots and stems ending in dentals show the regular change of *\*-d-t-* and *\*-t-t-* to *-ss-*, e.g., *passus* from the root *pat-*, *patior* ‘to suffer’, with subsequent simplification to *-s* if the syllable before the suffix was heavy, *\*kayd-to-* > *\*kaysso-* > *caesus* ‘cut’. **Oscan**, Umbrian, and other Sabellic languages attest numerous examples of this formation, e.g., Umbrian *screhto* ‘written’, Oscan *scrfīas* ‘written’ < \*skreyb<sup>h</sup>-; cf. Latin *scrīptus*.

The future active participle was a Latin development. In most cases, the suffix *-ūrus*, *-a*, *-um* was added to the stem of past participle, e.g., *dict-ūrus* ‘about to say’.

The PIE perfect active participle is not attested in Latin, but it may have survived in a few Sabellic formations, e.g., Oscan *sīpus* ‘knowing’ < \*sēpwōs, Volscian *sepu* < \*sēpwōd, though other explanations of these forms are possible.

The gerundive is an Italic innovation, but its origins remain uncertain. The suffix *-ndo-*/*-endo-* was added to the present stem in Latin, e.g., Latin *portandum* ‘to be carried’, *faciendum* ‘to be done’. In Oscan and Umbrian *-nd-* assimilated to *-nn*, e.g. Oscan **ūpsannūm** ‘to be done’, m. acc. sg., and Umbrian **pelsans** ‘to be buried’, m. nom. sg.

The supine in *-tum* is the accusative singular of a *tu*-stem verbal noun. It was found predominantly in construction with verbs of motion, e.g., Umbrian **avef anzeriatu etu** ‘go to observe the birds’, Old Latin *abiit ambulātum* ‘she went away to take a walk’. In Latin the supine in *-tū* is found after adjectives such as *facilis* ‘easy’, e.g., *facilis dictū* ‘easy to say’.

The present infinitives in Italic have roots in PIE nominal formations. The present infinitive in Latin, which ends in *-re*, is in origin an *s*-stem locative *\*-si*. The present passive infinitive of verbs belonging to conjugation 3 derive from root nouns inflected for the dative case, e.g., *capī* ‘to be seized’ < \*kapey. Present passive infinitives of conjugations 1, 2, and 4 are modeled on the active infinitive, e.g., *portāre* ‘to carry’ vs. *portārī* ‘to be carried’. An alternative passive ending *-ier*, which must go back to an instrumental case form augmented by the passive suffix *-r*, is attested in Old Latin and occasionally later, e.g., OL FIGIER ‘to be fashioned’. The present infinitive in Sabellic languages was an *o*-stem accusative, e.g., Oscan *ezum* ‘to be’, *moltaum* ‘to fine’. The present passive infinitive ends in *-fi* in Umbrian, e.g., *pihafi* ‘to be purified, which probably comes from an instrumental *\*-d<sup>h</sup>eyh<sub>1</sub>*. Oscan had the same suffix, but it was outfitted with the characteristic sign of the passive, e.g., **sakarafir** ‘to be consecrated’.

The Latin perfect active infinitive was modeled on the present. The ending *-se* was added to the perfect stem that ended in the formant *-is-*, e.g., *dixisse* (*dīks-is-se*). The corresponding passive is a periphrastic form, e.g., *dictum esse*. The system of infinitives

in Latin was filled out by future active and passive forms, e.g., active *dīcturum esse* ‘to be going to say’ and passive *dictum īrī* ‘to be going to say’. Corresponding forms are not attested in the Sabellic languages.

## SYNTAX

Nuclear sentences consist of a verb and its dependents. As in older IE languages, Latin and the other Italic languages marked the relationships between dependents and verbs by case endings. Grammatical subjects were inflected in the nominative, direct objects in the accusative, unless the lexical properties of the verb specified a different case, and indirect objects in the dative. Non-obligatory constituents, typically adverbial in nature, were expressed by a particular case or by a prepositional phrase, e.g., *noctū* ‘at night’, *decem diēs* ‘for ten days’, *cum exercitū suō* ‘with his own army’. The subject of the verb could be omitted if it was recoverable from context. Direct objects could be omitted as well, if context permitted. Agreement was found in two areas of the syntax: (1) verbs were inflected for the person and number of the subject; and (2) adjectival modifiers were inflected for the gender, number, and case of their head nouns. The inflection of dependents, omission of subjects, and agreement phenomena are at play in the following passage from Caesar’s *De Bello Gallico* (6.13.1–8). In (10c) the verb *concurrit* agrees in number, 3rd singular, with the subject *numerus*. In (10e) the adjective *summam* agrees in gender, number, and case with its head noun, *auctōritātem*. And again in (10e) the subject of *cōstituunt* is recoverable from the second clause in (10d).

- (10) (a) *In omnī Galliā eōrum hominum quī aliquō sunt numerō atque honōre genera sunt duo.* (b) *dē hīs duōbus generibus alterum est druidum, alterum equitum.* (c) *Illī rēbus diuinīs intersunt, sacrificia pūblica ac priuata procūrant, religiōnēs interpretantur.* (d) *ad hōs magnus adulescentium numerus disciplīnae causā concurrit, magnōque hī sunt apud eōs honōre.* (e) *nam fere de omnibus contrōuersīs publicīs priuatīsque cōstituunt.* (f) *hīs autem omnibus druidibus praeest ūnus, quī summam inter eōs habet auctōritātem.*

(a) ‘In all of Gaul, there are two classes of men who are of some rank and honor. (b) Of these two classes, one is Druid, the other knight. (c) The former are concerned with religious matters; they take care of public and private sacrifices; they explain religious phenomena. (d) A number of young men come to them [Druids] for training, and these [Druids] stand in great honor among them. (e) For they [Druids] make decisions about almost all public and private disputes. (f) One man stands at the head of all of these Druids, who has the greatest authority among them.’

### Constituent order

The unmarked arrangement of the major constituents in a sentence in Classical Latin is subject-object-verb (SOV), but that order it is not obligatory. SOV order is even more frequent in subordinate clauses where textual or pragmatic factors are less likely to have resulted in the movement of a constituent to the front or the end of the clause. In (10c), (10d), and (10e), the underlined verbs, apart from the copular verb *sunt*, occupy the final position in their clauses. In (10f) the verb *habet* (underlined) has been displaced from final position by the direct object *auctōritātem*, which has been moved to the end of the clause for emphasis, so-called right dislocation.

Classical Latin prose permits the distraction of constituents, usually resulting from the movement of elements to be emphasized either to the front or to the end of the sentence. In (10f) the noun *auctōritātem* has been moved out of the object-noun phrase (*summan auctōritātem*) and placed at the end of the sentence. In (10d) the constituents of the noun phrase *magnō honōre* were distracted and placed at the beginning and end of the clause.

Classical Latin poets did not adhere to the SOV standard as rigorously as prose writers for reasons that have to do in part with metrical necessity and in part with stylistics. The poets employ distraction more aggressively, and constituents at all levels of structure may be pulled apart for emphasis, focus, or artistic effect such as alliteration and assonance. In some instances the parts of constituents were “scrambled” with one another so that the continuity of every constituent in the sentence is interrupted. The following is from a poem by Horace. Distracted constituents are indexed by superscripts.

- (11) *mē tabulā<sup>b</sup> sacer<sup>a</sup> | uōtīuā<sup>b</sup> pariēs<sup>a</sup> indicat uuida<sup>c</sup> | suspendisse potentī<sup>d</sup> | uestimenta<sup>e</sup> maris deō<sup>d</sup>* (Hor. *Carm.* 1.5.13–16) ‘the sacred wall with (my) votive tablet proclaims that I have hung my soaked garments to the powerful god of the sea’.

The SOV arrangement favored by Latin prose authors is often thought to have preserved a stage of the language that was changing from “head final” to “head first”. This idea is supported by the fact that heads occur before modifiers in other constituents; that is to say, the unmarked order is Noun–Adjective (but not for “subjective” adjectives of the *bonus* type), Preposition–Noun Phrase, and Antecedent–Relative Clause. Furthermore, in the plays of Plautus, in main clauses, the order VO is statistically about as common as that of OV (Adams 1976). This view also accords well with the fact that in sub-literary texts of the first few centuries CE, VO order was the default. It is noteworthy that the variable order of constituents attested in Plautus contrasts with the order found in bureaucratic prose of roughly the same time period, which is predominantly SOV (Clackson & Horrocks 2007: 27–29). In the decree of the Roman Senate regarding participation in Bacchic cults, which was issued in 186 BCE, the order of the major constituents is consistently SOV. This suggests that SOV order was becoming, if it had not already become, one of the defining characteristics of elite Latin prose, thus distinguishing it from sub-elite varieties of the language.

The PIE word-order rule known as Wackernagel’s Law, by which enclitic elements were positioned after the first word or constituent of a sentence, is attested to some degree in Latin. Sentential particles, e.g., *enim*, *uērō*, *igitur*, and *autem*, as in (10f) above, regularly occupy the 2nd position in a Latin sentence. But it was more common in Latin for words and constituents that were highlighted or emphasized to serve as the hosts for enclitic elements such as 1st and 2nd person pronominal forms or the verb *esse* ‘be’. Since these highlighted constituents were often moved into the 1st position in their clause (or phrase), unaccented pronominal forms and unaccented forms of the verb ‘be’ generally came to occupy the 2nd position (Adams 1994a and 1994b).

The unmarked order of constituents in the Sabellic languages was probably also SOV, although one must point out that many of the inscriptions are short and formulaic, and modeled on epigraphic types inherited from Greek and Latin.

### Subordinate clauses

Latin poetry and prose, and Latin administrative prose, were home to a rich array of complex clause structures, some of which were elaborated throughout the Classical period

under the influence of comparable structures in Greek literature (Clackson & Horrocks 2007: 184–198). A short survey of these structures follows.

### Complementation

Sentential complements, that is to say, clauses that functioned as one of the arguments of a verb, had two main structures in Latin. The structure of the complement depended to a significant extent on the meaning of the verb. Some verbs governed a finite clause, others a non-finite clause containing an infinitive and an accusative subject.

1. The Accusative + Infinitive construction was the rule after verbs of saying (*dīcere*), thinking (*putāre*, *arbitrārī*), knowing (*scīre*, *cognōscere*), understanding (*intelligere*), and perceiving (*audīre* ‘hear’, *uidēre* ‘see’, *sentīre* ‘perceive’). The accusative was the subject of the infinitive and was rarely omitted from surface structure, even when it was co-referential with the subject of the main clause, as it is in (12).

(12) *dīcēbās quondam [solum tē nōsse Catullum] | Lesbia* (Catul. 72.1–2)  
 ‘You used to say [that you knew only Catullus], Lesbia.’

In the passive counterpart of this construction, cited in (13), the underlying subject of the infinitive *Epamīnōndās* has been moved out of the Accusative + Infinitive clause to be the subject of the main clause and for this reason is inflected in the nominative. In such sentences modifiers of the subject were also inflected in the nominative.

(13) *Epamīnōndās [fidibus praeclarē cecinisse] dīcitur* (Cic. *Tusc.* 1.4)  
 ‘Epaminondas is said [to have played brilliantly on the lyre].’

2. After verbs and verb phrases of “emotion”, e.g., *gaudeō* ‘rejoice’, *indoleō* ‘regret’, or *molestē ferō* ‘find annoying’, two structures were permitted: Accusative + Infinitive and *quod* ‘that’ + indicative.
3. Interrogative particles (*num* ‘whether’, *-ne* ‘if’), adverbs (*ubi* ‘where’), and pronominal forms (*quis*, *quid* ‘who, what’) introduced indirect questions. The verb was in the subjunctive mood.

(14) *senex dīcitur quaesisse num illud carmen dēsipientis uidērētur* (Cic. *Sen.* 22)  
 ‘The old fellow is said to have asked whether this seemed the poem of a man out of his mind.’

4. Indirect commands were introduced by verbs of ordering (*iubēre*, *imperāre*), forbidding (*uetāre*), permitting (*sinere*), exhorting (*hortārī*), requesting (*rogāre*), resolving (*dēcernere*), and warning (*monēre*). Most verbs took subordinate clauses of the finite type. Such clauses were introduced by *ut(ī)/nē*, and the verb was in the subjunctive mood. The verbs *iubēre* ‘order’ and *uetāre* ‘forbid’, however, took the Accusative + Infinitive construction. Compare sentences (15) and (16).

(15) *hortātur eōs [nē animō deficiant]* (Caes. *Civ.* 1.19)  
 ‘He encourages them [not to lose heart].’

(16) [*Labiēnum iugum montis ascendere*] *iubet* (Caes. *Gal.* 1.21)  
 ‘He orders [Labiēnus to ascend the ridge of the hill].’

5. Complement clauses after verbs of fearing have a structure similar to indirect commands, as do complements following verbs of hindering, though the conjunctions that introduced them were distinct. The conjunctions *nē* / *nē nōn* introduced complement clauses after verbs of fearing (*timēre*, *uerērī*), and *quīn* introduced complement clauses after verbs of hindering (*impedīre* ‘impede, obstruct’, *prohibēre* ‘prevent, forbid’).

The range of complement clause types was more limited in sub-elite Latin of the imperial period. The number of verbs that took the Accusative + Infinitive construction decreased; complements introduced by *quod* or *quia* extended the contexts in which they appeared and eventually become the norm. The Accusative + Infinitive construction survives only after verbs of perception in Romance languages.

### Adverbial clauses

Other subordinate clause types functioned as adverbial modifiers. These clause types covered a range of functions, including time, cause, purpose, concession, comparison, and result. Adverbial clauses in Latin were, with one exception to be discussed below, finite clauses.

1. Purpose clauses were introduced by the conjunction *ut* ‘so that’. The verb was in the subjunctive mood. If the clause was negated, the conjunction was *nē*. Result clauses were similar in structure to purpose clauses, but if the clause was negated, it was introduced by *ut nōn*. Even if the result clause referred to an actual event, the mood of the verb was subjunctive.
2. Temporal clauses were introduced by the following conjunctions: *cum*, *ut*, *ubi* ‘when’; *dum* ‘while’; *dum*, *dōnec*, *quoad* ‘as long as’; *dum*, *dōnec*, *quoad* ‘until’; *postquam* ‘after’; and *antequam*, *priusquam* ‘before’. The mood of the verb was indicative in most cases, but in clauses that referred to events that were anticipated or generic, the subjunctive was making inroads in Classical Latin.
3. Causal clauses introduced by *quod*, *quia*, *quoniam*, and *quandō*, ‘because, since’ generally took a verb in the indicative mood, but the conjunction *cum* required a subjunctive.
4. Conditional clauses were introduced by *sī* ‘if’. If the relationship between the verbs in the *sī*-clause and the main clause was real, the verbs were in the indicative mood. If the relationship between the verbs in the clauses was considered possible or was counterfactual, the verbs were in the subjunctive mood.
5. The ablative absolute construction consisted of a noun phrase as subject and a participle as predicate, and both were inflected in the ablative case. Nominal structures were also possible, in which case the subject phrase was followed by a predicative adjective or noun. The absolute was in origin a circumstantial clause that stood in a variety of adverbial relationships to the main verb. Ablative absolutes were rarely introduced by conjunctions, so the subordinate relationship was determined by the meaning of the predicate and the sentential context. The participle expressed relative time.

Prototypical absolute constructions were short, consisting of a noun and a participle (17). In Old Latin this was almost always the case, but during the Classical period, at the hands of authors who were influenced by comparable structures in Greek, this clause

type became longer and more complex in its structure. Imperial prose writers such as the historian Tacitus sometimes specified the subordinate role of the absolute by adding a subordinating conjunction (18).

- (17) *Caesar [acceptīs litterīs] nuntium mittit* (Caes. *Gal.* 5.46)  
 ‘[After the letters were received], Caesar sends a messenger.’
- (18) *aliī [quamuis effugiō patente] interiēre* (Tac. *Ann.* 15.38.7)  
 ‘Some died, although an escape route was open.’

### Relative clauses

In Classical Latin relative clauses typically followed the antecedent noun phrase, particularly when they were restrictive, but other arrangements were possible. Relative clauses were sometimes placed inside the phrase containing the antecedent, as in (19), or placed in initial position for reasons having to do with style, as in 20, which reflects an order common in legal texts. Antecedent phrases and relative pronouns are marked in bold.

- (19) *ab iīs [quōs miserat] explorātōribus Ø et ab Rēmīs cognouit . . .* (Caes. *Gal.* 2.5.4)  
 ‘he learned from those scouts that he had sent and from the Remi . . .’
- (20) [**quibus ita est interdictum**] **hī** *numerō impiōrum ac scelerātōrum habentur* (Caes. *Gal.* 6.13.7)  
 ‘[to whom it has been banished in this manner], these (folks) are considered in the number of the impious and the wicked.’

Other relative clause structures are found in Classical Latin. The correlative type, exemplified in (21) and (22), has comparable structures within Italic and more widely within Indo-European. In this type the relative clause occupied the 1st position in the sentence, and the antecedent was incorporated into the relative clause. The antecedent was then referenced in the main clause by a pronominal form, as in (21), or by the antecedent and its pronominal modifier, as in (22).

- (21) [*quae pars cūiūtātis calamitātem populō Rōmānō intulerat*] **ea** *prīnceps poenās persoluit* (Caes. *Gal.* 1.12)  
 ‘[Which part of the state had brought disaster on the Roman people], that (part) first paid the price.’
- (22) [**QUEM AGRUM EOS UENDERE HEREDEMQUE SEQUI LICET**] **IS** *AGER UECTIGAL NEI SIET* (CIL I<sup>2</sup>.584.5)  
 ‘Which property it is permitted for them to sell and to follow an heir, that property shall not be subject to taxation.’

Relative pronouns were sometimes subject to a syntactic process known as “attraction”. Relative pronouns could be attracted into the case of their antecedents, as in (23), where the accusative object of *cōnsuēstī* is attracted to the genitive case of *eōrum*. In (24) the antecedent *urbem* was attracted into the case of relative pronoun *quam* (*attractio inuersa*



‘inverse attraction’). Inverse attraction is well attested in Old Latin, but is rare in Classical Latin. As a result, Classical Latin practice could well be an archaism.

- (23) *sī aliquid agās eōrum [quōrum cōnsuētī]* (Cic. *Fam.* 5.14)

‘If you should do any of the things that you were accustomed to do.’

- (24) *urbem [quam statuō] uestra est* (Virg. *Aen.* 1.573)

‘The city [that I am founding] is yours.’

### Sequence of tenses

In indirect questions and indirect commands the tense of the verbs in the subordinate clauses was governed by the tense of the main verb. Verbs that were present or future governed a subordinate verb that was present or perfect; verbs that were past tense (imperfect, perfect, pluperfect) governed a subordinate verb that was imperfect or pluperfect, as in (25). Note that the tense of the verb in the original question in (25) was perfect tense (‘Why did you all come to my house?’). This syntactic rule, which is also known as back shifting, is attested in the Sabellic languages.

- (25) *(Lentulus) quaesīuit ā Gallīs quam ob rem domum suam uēnissent* (Cic. *Cat.* 3.11)

‘He (Lentulus) asked the Gauls why they had come to his house.’

### Other subordinating structures

Subordinate ideas could also be conveyed by non-finite structures, including participles, gerundives, and supines. A selection of the possibilities is discussed below.

Purpose could be indicated by a supine when in construction with verbs meaning ‘come’ and ‘go’ (26). After the preposition *ad*, purpose could be indicated by a gerundive phrase that agreed with its head noun in gender, number, and case (27). Gerundives indicating subordinate ideas were the rule after verbs such as *cūrāre* ‘to take care (that such and such happens)’ (28).

- (26) *legātōs ad Caesarum mittunt [rogātum auxilium]* (Caes. *Gal.* 1.11.3)

‘They send legates to Caesar to ask for assistance.’

- (27) *[ad bella suscipienda] Gallōrum alacer est animus* (Caes. *Gal.* 3.19.6)

‘The character of the Gauls is keen to undertake wars.’

- (28) *[pontem in Ararī faciendum] cūrat* (Caes. *Gal.* 1.13.1)

‘He takes care to build a bridge over the Arar.’

Participles serving as attributes of noun phrases conveyed a wide range of subordinate ideas. In (29) the phrase *agrūm Gallicūm dīuidentī* had the function of a temporal clause introduced by ‘while’. In (30) the bracketed phrase, which is an attribute of *Dionysius*, had a causal function.

- (29) *C. Flaminiō restitit [agrūm Gallicūm dīuidentī]* (Cic. *Sen.* 11)

‘He resisted Gaius Flaminius [while he [Flaminius] was apportioning Gallic territory].’

- (30) *Dionysius [cultrōs metuēns tōnsōrēs] candente carbōne sibi adūrēbat capillum*  
(Cic. *Off.* 2.25)

‘Dionysius, because he was afraid of the barber’s shears, used to singe his hair with a piece of white-hot coal.’

## LEXICON AND WORD FORMATION

From a historical perspective the Latin lexicon is made up of strata of vocabulary, the deepest of which can be traced back to the parent language. Latin preserved IE vocabulary in a wide range of semantic spheres, e.g., kinship terms (*pater* ‘father’, *māter* ‘mother’, *frāter* ‘brother’, *soror* ‘sister’), basic numbers (*duo* ‘two’, *trēs* ‘three’, *quattuor* ‘four’, *quinque* ‘five’, etc.), body parts (*pēs*, *pedis* ‘foot’, *manus* ‘hand’), domesticated animals (*sus* ‘pig’, *ouis* ‘sheep’), weather terms (*nix*, *niuis* ‘snow’, *pluit* ‘it’s raining’), and so forth.

A few words, primarily of a religious nature, can be assigned to the Proto-Italic period, e.g., Latin *sacer* ‘sacred’, nom. sg. m., Umbrian **sacru** ‘sacred’, nom. sg. n.; Latin *sacrem* ‘sacrificial victim’, acc. sg., Umbrian **sakre**, acc. sg., Oscan **sakrim**, acc. sg.; Latin *sanctum* ‘inviolable’, Oscan **sahtúm**; and Latin *pius* ‘devout’, Oscan **pihiúí**, dat. sg. A smaller number of words belong to the secular sphere, e.g., Latin *ūtī* ‘to use’, Oscan **úitiúf** ‘usage’, nom. sg.; Latin *cēna* ‘meal, banquet’, Oscan **kersnu** ‘banquet’, Umbrian *çesna* ‘meal’. In addition, a few words that are IE in origin underwent semantic developments peculiar to Italic. PIE \*deyk- ‘to point out, show’, for example, acquired the meaning ‘say’ in Latin and Oscan.

Divergences existed in the Italic lexicon as well. In Latin the words *filius* ‘son’ and *filia* ‘daughter’ replaced the PIE words for ‘son’ and ‘daughter’, but they were preserved in the Sabellic languages, e.g., Oscan **pukluí**, ‘son’, m. dat. sg., South Picene **puqloh**, m. dat. sg. < \*putlo-; Oscan **futír** ‘daughter’, f. nom. sg., **futrei**, f. dat. sg. < \*d<sup>h</sup>ugh<sub>2</sub>tēr.

Borrowings in Latin cut across the chronological periods of the language. A few words have come from a Sabellic language of central Italy, e.g., *bōs* ‘cow’, *rufus* ‘red’, and *popina* ‘cook shop’. Further afield, Celtic incursions into the peninsula beginning in the 6th century BCE provided a small number of words, e.g., *carrus* ‘wagon’. In the 3rd century BCE, as contact with Greek speakers intensified, Greek loans began to enter the language in greater numbers. The plays of the Roman comedian Plautus are filled with Greek loans in semantic spheres ranging from business and finance to medicine and education. Among the Roman elite who took up the study of philosophy, oratory, literature, and philology, Greek was the source of many technical terms. The words were borrowed with varying degrees of accommodation to the Latin phonological and morphological systems, depending on the date of the borrowing, on the education of the borrower, and on the reason for the borrowing. For some Greek technical terms, new Latin words were coined rather than borrowed, e.g., *essentia* ‘substance’ for Greek οὐσία; for others, existing Latin words were semantically extended to cover Greek concepts, e.g., *cāsus* ‘accidence’ for Greek πῶσις.

Throughout the history of Latin, productive morphological processes augmented the lexical inventory. Abstract nouns and result nouns were derived from verb roots and verb stems by suffixation of *-men* (*fragmen*, *fragminis* ‘fragment’), *-mentum* (*fragmentum*, *fragmenti* ‘fragment’), *-iōn-* (*occidiō*, *occidiōnis* ‘slaughter’), and *-tiōn-* (*aedificatiō*, *aedificatiōnis* ‘construction’). The suffixes *-(i)tāt-* (*bonitās*, *bonitātis* ‘goodness’) and *-(i)tūt-* (*uirtūs*, *uirtūtis* ‘manliness’) also formed abstracts from adjective stems. Agent nouns were made with the suffix *-tōr*, which was added to verb roots and verb stems, e.g., *amātor*, *amātōris* ‘lover’, *scrīptor*, *scrīptōris* ‘writer’. The adjective suffix *-ānus*,

which originated as a re-segmentation of adjectives derived from 1st declension nouns, was very productive, e.g., 3rd declension *urbānus* ‘of the city’, 2nd declension *mundānus* ‘of the world’, on the model of 1st declension *Romānus* ‘of Rome’. The verb suffix *-ā* was extracted from 1st conjugation verbs and was used to form denominatives from consonant-stems, e.g., *nomināre* ‘to call by name’, built on the stem *nomin-* ‘name’. The suffix *-(i)tā*, which originally formed iterative/intensives, eventually replaced the corresponding simple formations, e.g., *cantāre* ‘to sing’ to *canere*, *iactāre* ‘to throw’ to *iacere*, *dictāre* ‘to say’ to *dīcere*. Compound verbs were formed by addition of prefixes to verb stems, e.g., *dūcere* ‘to lead’, *ēdūcere* ‘to lead out’. Compound noun formations were not as productive in Latin as they were in IE languages like Greek, but compounds such as *agricola* ‘farmer’, *auceps* ‘bird catcher’, *aurifex* ‘goldsmith’, and *iūdex* ‘judge’ may point back to a more productive period. In classical literature, compounds began to reappear in greater numbers, but they were created on models from Greek literature, e.g., *armiger* ‘bearing arms’.

### FURTHER READING

Editions of Latin literary texts are published in numerous text series. The Oxford Classical Texts series is the standard. Latin epigraphic materials are collected in the *Corpus Inscriptionum Latinarum*. Rix (2002) is an *editio minor* of Sabellic inscriptions. Venetic inscriptions published before 1987 are in Prosdocimi 1987. Venetic inscriptions recovered after 1987 are published annually in *Studi Etruschi*, in the section on Italic epigraphy (*Rivista di Epigrafia Italica*).

Clackson 2011 covers the Latin language from all angles. Rosen 1999 focuses on the development of literary Latin. Historical Latin syntax is treated in the multi-volume work edited by Baldi and Cuzzolin (2009). Allen’s *Vox Latina* (1978<sup>2</sup>) is the basic introduction to Latin phonology. A number of excellent reference works in the comparative/historical tradition have been published in the last 20 years. Baldi 1999, Clackson and Horrocks 2007, Meiser 1998, Sihler 1995, and Weiss 2009 are particularly useful. For the prehistory of the Latin verb, see Meiser 2003. For discussion of VOL and OL inscriptions Vine 1993 and Wachter 1987 are essential. The following etymological dictionaries cover Latin and the other Italic languages: Ernout and Meillet 1985, Untermann 2000, and de Vaan 2008. For Latin dialects from regional and social perspectives, see Adams 2007 and 2013. Bilingualism in the Roman world is the topic of Adams 2003.

Fortson 2010, chapter 13, is a concise and informative introduction to Italic and the Italic languages (Venetic is discussed in Chapter 20, “Fragmentary Languages”). Wallace 2007 is an overview of the Sabellic languages.

### REFERENCES

- Adams, J.N. 1976, “A typological approach to Latin word order,” *Indogermanische Forschungen* 81: 70–99
- Adams, J.N. 1994, *Wackernagel’s Law and the placement of the copula esse in classical Latin*, *PCPhS Supplement* 18, Cambridge Philological Society, Cambridge
- Adams, J.N. 1994b, “Wackernagel’s Law and the position of unstressed personal pronouns in Classical Latin,” *Transactions of the Philosophical Society* 92: 103–178
- Adams, J.N. 2003, *Bilingualism and the Latin Language*, Cambridge University Press, Cambridge

- Adams, J.N. 2007, *The Regional Diversification of Latin 200 BC – AD 600*. Cambridge University Press, Cambridge
- Adams, J.N. 2013, *Social Variation and the Latin Language*, Cambridge University Press, Cambridge
- Allen, W. Sidney 1978<sup>2</sup>, *Vox Latina. The Pronunciation of Classical Latin*, Cambridge University Press, Cambridge
- Baldi, Philip 1999, *The Foundations of Latin*, Mouton de Gruyter, New York–Berlin
- Baldi, Philip & Cuzzolin, Pierluigi 2009, *New Perspectives on Historical Latin Syntax. Volume 1: Syntax of the Sentence*, Mouton de Gruyter, New York–Berlin
- Clackson, James (ed.) 2011, *A Companion to the Latin Language*, Wiley-Blackwell, Oxford–Malden, MA
- Clackson, James & Horrocks, Geoffrey 2007, *The Blackwell History of the Latin Language*, Blackwell, Oxford
- Corpus Inscriptionum Latinarum*, Berlin, 1862–2004
- de Vaan, Michel 2008, *Etymological Dictionary of Latin and the Italic Languages*, Brill, Leiden–Boston
- Ernout, Alfred & Meillet, Antoine 1985, *Dictionnaire étymologique de la langue latine*, Klincksieck, Paris
- Fortson, Benjamin W. 2010<sup>2</sup>, “Italic,” in: *Indo-European Language and Culture. An Introduction*, Wiley-Blackwell, Oxford–Malden, MA: 274–307
- Holloway, R. Ross. 1994, *The Archaeology of Early Rome and Latium*, Routledge, London–New York
- Meiser, Gerhard 1998, *Historische Laut- und Formenlehre der lateinischen Sprache*, Wissenschaftliche Buchgesellschaft, Darmstadt
- Meiser, Gerhard 2003, *Veni, vidi, vici: Die Vorgeschichte des lateinischen Perfektsystems*, Beck, Munich
- Prodocimi, Aldo L. 1987, “La lingua,” in: Fogolari, G. & Prodocimi, Aldo L. (eds.), *I Veneti antichi: Lingua e cultura*, Studio Editoriale Programma, Padova: 223–440
- Rix, Helmut 1994, “Latein und Sabellisch: Stammbaum und/oder Sprachbund,” *Incontri Linguistici* 17: 13–29
- Rix, Helmut 2002, *Sabellische Texte. Die Texte des Oskischen, Umbrischen, und Südpikenisches*, Carl Winter, Heidelberg
- Rosen, Hannah 1999, *Latine loqui: Trends and Directions in the Crystallization of Classical Latin*, W. Fink, Munich
- Sihler, Andrew L. 1995, *New Comparative Grammar of Greek and Latin*, Oxford University Press, Oxford
- Stuart-Smith, Jane 2004, *Phonetics and Phonology: Sound Change in Italic*, Oxford University Press, Oxford
- Untermann, Jürgen 2000, *Wörterbuch des Oskisch-Umbrischen*, Carl Winter, Heidelberg
- Vine, Brent 1993, *Studies in Archaic Latin Inscriptions*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck
- Wachter, Rudolf 1987, *Altlateinische Inschriften*, Peter Lang, Bern
- Wallace, Rex 2007, *The Sabellian Languages*, Lincom Europa, Munich
- Weiss, Michael 2009, *Outline of the Historical and Comparative Grammar of Latin*, Beech Stave Press, Ann Arbor–New York

# CELTIC

*Patrick Sims-Williams*

## INTRODUCTION

The surviving Celtic languages fall into two groups: (a) the *Brythonic* (or Brittonic or British) group (Ternes 2011); and (b) the *Gaelic* (or Goidelic or Irish) group. The two are very distinct and have been mutually unintelligible for well over a millennium. To the Brythonic group belong Welsh, spoken widely in Wales, and Breton, spoken in the west of Brittany. Cornish, the language of Cornwall, which was very similar to Breton, died out as a natural language in the eighteenth century. To the Gaelic group belong Irish (or Irish Gaelic), spoken mainly in the west of Ireland, and Scottish Gaelic, spoken mainly in the west of Scotland. Manx, the Gaelic language of the Isle of Man, died out as a natural language in the twentieth century. Celtic languages are also spoken in the Americas, as a result of the modern diaspora of Celtic-speaking peoples, so that, for example, there are Welsh/Spanish bilinguals in Patagonia in Argentina, and Scottish Gaelic/English bilinguals in Nova Scotia in Canada.

All these surviving Celtic languages (including Breton!) are known collectively as *Insular Celtic* languages, as opposed to the ancient *Continental Celtic* languages, e.g. Gaulish, Galatian, Celtiberian, etc., which were all dead by AD 500 and mostly much earlier. The term “Insular” refers to the two islands of Ireland and Britain. From these the Gaelic and Brythonic languages spread: in about the fifth century AD Scotland was settled by emigrants from Ireland (the *Scotti*), while Armorica was settled by emigrants from southern Britain, becoming known as Brittany (Breton *Breiz* < \*Brittia). The theory that the Breton language includes a substratum of indigenous Armorican Celtic is unproven but not impossible.

The Brythonic dialects began to diverge into West British (> Welsh) and South-West British (> Cornish and Breton) in about the fifth century AD, but probably remained mutually intelligible for several centuries. The Gaelic dialects began to diverge in about the tenth century AD, but there was a common literary language to the end of the Middle Ages, and even today Irish and Scottish Gaelic are much more similar than are Breton and Welsh, which have long been mutually unintelligible. The geographical reason is obvious: there has always been easy travel between Scotland and northern Ireland. Another factor is that Welsh has been in contact with English, whereas Breton has been in contact with French. These contacts have influenced the syntax, morphology and phonology of the languages; for example, there are nasalized vowels in Breton and French, but not in Welsh and English.

Within most individual surviving languages there are marked dialectal differences, so that Breton speakers from north and south may find it easier to communicate in French, while Gaelic speakers from north and south may prefer to talk English together. Nearly all adult Celtic speakers are bilingual.

A general impression of the divergence of the Celtic languages may be gained from comparing the ordinal numerals in Gaulish (McCone 1994: 208 = 2005: 300, Lambert 2002: 105–109, Hamp 2012) with those in Old Irish and Middle Welsh; see Table 7.1.

TABLE 7.1 COMPARISON OF THE ORDINAL NUMERALS

	Gaulish (first century AD)	Old Irish (eighth century)	Middle Welsh (thirteenth century)
1	<i>cintux(?)mos</i>	<i>cétn(a)e</i>	<i>kyn taf</i>
2	<i>al(l)os</i>	<i>tán(a)ise, aile</i>	<i>eil</i>
3	<i>trito(s), tr(itios)</i>	<i>tris</i>	<i>trydyd, f. tryded</i>
4	<i>petuarios</i>	<i>cethramad</i>	<i>pedwryyd, f. pedwarded</i>
5	<i>pinpetos</i>	<i>cóiced</i>	<i>pymhet</i>
6	<i>suexos</i>	<i>se(i)ssed</i>	<i>chwechet</i>
7	<i>sextametos</i>	<i>sechtmad</i>	<i>seithvet</i>
8	<i>oxtumetos</i>	<i>ochtmad</i>	<i>wythvet</i>
9	<i>namet(os)</i>	<i>nómad</i>	<i>nawvet</i>
10	<i>decametos</i>	<i>dechmad</i>	<i>decvet</i>

**Celtic**

The term “Celtic”, as applied to the Insular Celtic-speaking peoples and their languages, is a modern one. These peoples did not refer to themselves and their languages as “Celtic” until recently. For example, the medieval Irishmen were *Goidil* (an opprobrious name derived from Brythonic, cf. W *grŷydd* ‘wild’), and their language was *Góidelach*, and medieval Welshmen regarded themselves as *Brython* (< Lat. *Brittones*) or *Cymry* (< \*kom-brogī ‘co-countrymen’) and their language was *Cymraeg*. (The Welsh themselves do not use the English name *Welsh* < OEng. *w(e)alh* ‘foreigner, or slave, mostly speaking a Romance or Gallo-Brythonic language’. This may derive ultimately from the Continental ethnic name *Volcae*.) There is no evidence that medieval Brythonic and Gaelic speakers recognized their special linguistic kinship; this was discovered by early comparative philologists such as the Scot George Buchanan (in his *Rerum Scoticarum Historia*, 1582) and the Welshman Edward Lhuyd (in his *Archaeologia Britannica*, 1707). The modern Romantic idea of a pan-Celtic ethnic unity and “Celtic national character” has had some influence even in the Celtic-speaking countries, but really derives from foreign works such as Ernest Renan’s *La Poésie des races celtiques* (1854) and Matthew Arnold’s *The Study of Celtic Literature* (1866).

The linguistic term “Celtic” derives from the usage of ancient and early medieval Greek and Latin writers, who only use it of *Continental* Celtic languages; for example, as late as the ninth century, Heiric of Auxerre, *Vita S. Germani* I.353, explains the place-name *Augustidunum* as meaning ‘Augusti mons’ in *Celtica lingua*. *Celtica lingua*, in fact, seems to have been equivalent to *Gallica lingua*, the term used, for example, in the sixth century by Venantius Fortunatus, who explained the old Gaulish place-name *Vernemetis* as “fanum ingens” (‘great temple’) (*Carmina* I.ix.9–10). Because close ethnic and linguistic similarities between Gaul and Britain had been noted by writers from Tacitus (*Agricola*, 11) down to the Renaissance, it seemed reasonable to early modern scholars to apply the term “Celtic” to Brythonic as well. Thence it was extended to the languages of Ireland and Scotland, on comparative philological grounds, even though the Insular Celtic languages may never have been called “Celtic” in Antiquity (Sims-Williams 1998; cf. Blom 2009).

Despite its dubious origin, the term “Celtic” remains a useful label for a distinct family of IE languages. The phonological and lexical similarities between Gaulish, Brythonic and Gaelic are illustrated by the above forms *Augustidunum* (i.e. Gaul. \*-dūnon) and *Vernemetis* (i.e. Gaul. \*Wer-nemeton). With the first compare OIr. *dún* ‘fort’, OBret. *din* gl. ‘arx’, Corn. *dyn*, OW *din* (note /u:/ > /ɤ:/ > /i:/ in Brythonic). These Celtic cognates are distinct from the Germanic cognates such as OEng. *tūn* (> Eng. *town*) in not showing the effect of Grimm’s Law (/d/ > /t/; see p. 20, 392–393). With the intensive prefix *Ver-* ‘ingens’ compare OIr. *for* < \*wor- < \*wer- and OWCB *guor-* < \*wor-

< \*wer- (**Subdivision**, p. 358). These cognate forms differ from IE cognates such as Gr. ὕπερ in showing the well-known Common Celtic loss of IE /p/, i.e. \*uper > \*wer-. Lastly, Gaul. *nemeton* ‘fanum’ is probably related to Lat. *nemus*, Gr. νέμος ‘grove’, but the dental formation is only paralleled within Celtic, for example, OIr. *nemed* [nʲeμəd] (where μ represents a nasal bilabial fricative and cf. p. 367) glossing *sacellum* ‘sanctuary’ < \*neμeþan < \*nemeton, and Old Breton personal names in *-nemet* [neμed] and Old Welsh ones in *-nimet* [nəmed], both < British \*nemeton. This formation is already seen in the Celtic name \*Nemetios in an Etruscan inscription of the fifth century BC at Genoa: *MI NEMETIES* ‘I am [the tomb] of Nemetios’ (de Simone 1980).

## Origins

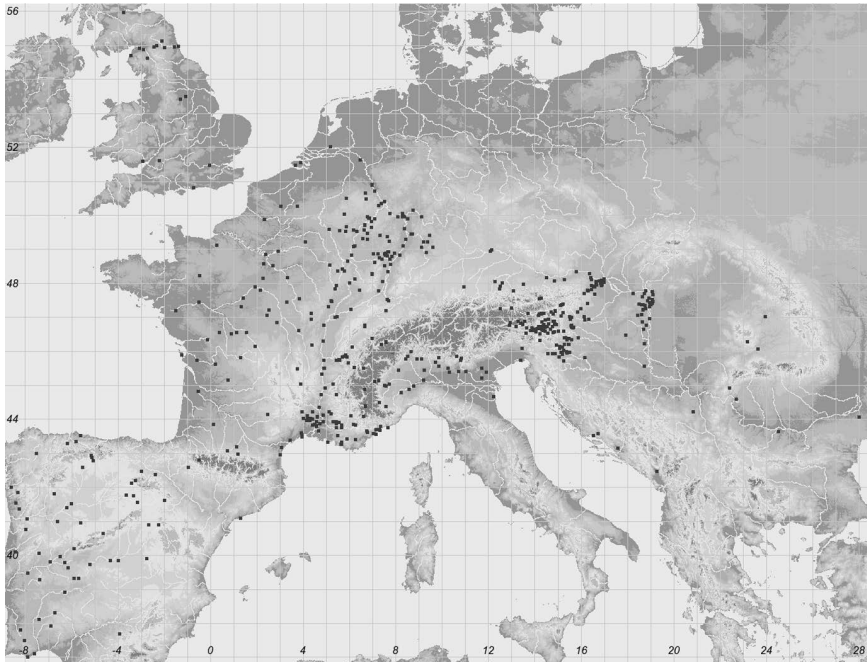
The original “homeland” of the Celtic speakers is unknown – but in any case the simplistic concept of a “homeland” is of limited validity, both in terms of ethnogenesis and in terms of language origins: a people may comprise diverse ethnic elements, and a language may derive from various sources (e.g. the English are ethnically both Celtic and Germanic, and the IE element in English comes via Latin and French as well as from Germanic). There are three main approaches to the problem.

Archaeological approaches have failed to locate the first speakers of Celtic. Archaeologists used to associate Celtic speech with either the early, so-called Hallstatt Iron Age culture or the later, so-called La Tène, Iron Age culture, or with both (see the maps in Powell 1980: 48, 115; cf. Mallory 1989: 96–107). While it is likely that Celtic speakers were found in these Hallstatt/La Tène areas, as in later centuries, we can rarely equate language with material remains: a correlation between La Tène finds and Celtic speech may work quite well in eastern Europe (Falileyev 2014), but it does not work in Spain or northern Italy, where Celtic inscriptions fail to correlate with La Tène material. More recent archaeologists, dissociating Celtic from the above Iron Age cultures, have speculated that a form of IE was already spoken in north-western Europe by 4000 BC and gradually developed into Celtic *in situ* (this phase began with Renfrew 1987: 249; cf. Meid 1989, Mallory 1989: 274). This extreme view is possible only *provided that* (it is a massive proviso) we can believe that communications over the four millennia BC would have allowed many parallel developments to take place in the Late IE and Early Celtic dialects. The fact is that we can only speculate about when Indo-European and then Celtic were first spoken on the Atlantic seaboard, including Britain and Ireland. The latest possible date for Celtic reaching Ireland is the first century AD (just before the Celtic names attested in Ptolemy’s *Geography*), and this has indeed been proposed as near the actual date (Schrijver 2014: 81). Archaeology can neither disprove nor confirm this proposal, and so far no light has been cast on these problems by ancient (or modern) DNA (Sims-Williams 2012a).

A second approach to the problem is to examine the references to *Keltoi*, *Celtae*, *Galatae*, etc. by Greek and Latin writers. There are several problems here: (a) These ethnic labels, which are of uncertain etymology (Sims-Williams 2011), may not always correspond to our modern linguistic use of the term *Celtic* (see p. 353 above). (b) Mediterranean observers may not have distinguished clearly between barbarian peoples such as *Celtae* and *Germani*, and probably simplified the true situation, especially at first (just as even today the French tend to apply the term *Anglais* to all the inhabitants of Britain). (c) There are very few references to “Celts” before 400 BC, the period when their violent expansion into Italy, the Balkans, Greece and Turkey began to bring them to the serious attention of Mediterranean writers. Earlier Greek writers associate them variously with Spain (the Marseilles *periplus* in the sixth century BC, if correctly reported in Avienus’ *Ora maritima*), with the hinterland of Marseilles (*Massilia*) (Hecataeus of Miletus, c. 500), and with the far west of Europe,

including the source of the Danube (!) in *Pyrene* = (?) Pyrenees (Herodotus, c. 450). (See Powell 1980: 11–15, Schmidt 1986a: 15, Tovar 1986: 79–80, Sims-Williams 2016.) Obviously these Greek writers were chiefly familiar with the Celts nearest to them, but there may well have been peoples calling themselves “Celts”, or who were called “Celts” by others, in northern Europe too. As we cannot prove this, however, early ethnographical testimonies are of limited help in defining the “Celtic” area of Europe before 400 BC. It used to be thought significant that Avienus’ names for the inhabitants of Ireland (*Hierni*, see **Consonantism**, p. 353) and Britain (*Albiones* cf. MW *elvyd* ‘world’ < \*albyo-, cognate with Lat. *albus* ‘white’, etc.) had specifically Celtic etymologies. We cannot be sure that these names originated in the British Isles, however, and anyway their linguistic Celticity is uncertain (Sims-Williams 1998: 20–21, Schrijver 2014: 78).

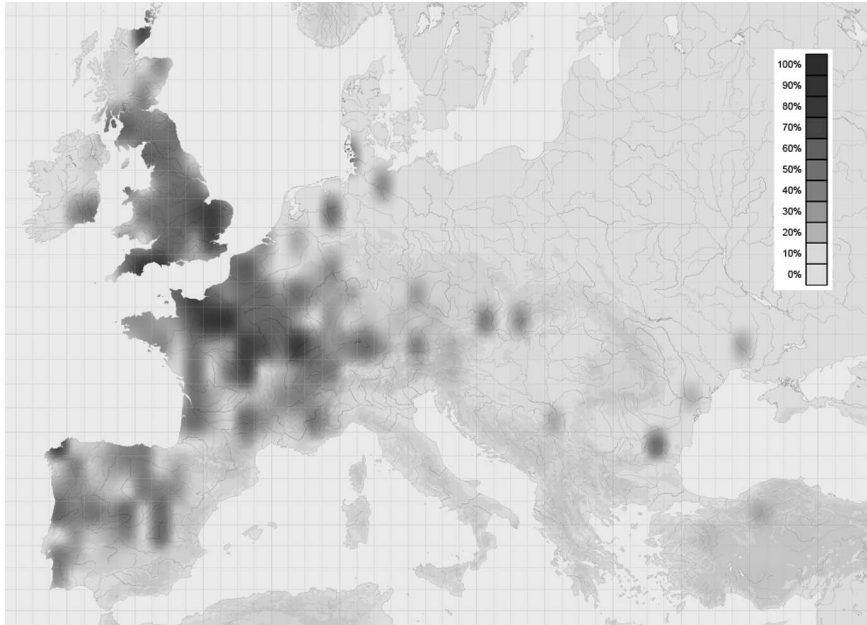
The third approach to the problem is to examine the distribution of apparently Celtic personal, place- and ethnic names (Sims-Williams 2007a, Raybould & Sims-Williams 2009, Falileyev et al. 2010) and the distribution of Celtic-language inscriptions, coin legends, graffiti and so on. Here, too, there are serious obstacles: (a) such data are of course chiefly recorded in the vicinity of Mediterranean cultures, where writing systems developed (on the orthographies of the Celtic languages see Russell 1995: 197–230 and de Hoz 2007); and (b) they mostly belong to the period shortly before the Christian era or still later, too late, that is, to shed light on the areas in which Celtic speech originated.



**MAP 7.1** LATIN PROVINCIAL LAPIDARY INSCRIPTIONS BEARING CELTIC PERSONAL NAMES

Source: Marilynne E. Raybould and Patrick Sims-Williams, *Introduction and Supplement to the Corpus of Latin Inscriptions of the Roman Empire Containing Celtic Personal Names*, CMCS Publications, Aberystwyth, 2009, map 2 (sample = compound names, the most recognizable type of Celtic name)





**MAP 7.2** DENSITY OF CELTIC-LOOKING ANCIENT PLACE-NAMES

Source: Marilynne E. Raybould and Patrick Sims-Williams, *Introduction and Supplement to the Corpus of Latin Inscriptions of the Roman Empire Containing Celtic Personal Names*, CMCS Publications, Aberystwyth, 2009, map 16 (created by M. Crampin on the basis of data in Sims-Williams 2007a: 164–165)

### Position within IE

The position of Celtic among the IE languages was approached in various ways during the twentieth century (cf. Schmidt 1979: 197). (1) Pokorny, Wagner and others emphasized Celtic's divergence from the IE model and explored the possibilities of substratum or "areal" influences from non-IE languages, for example, Semitic and Hamitic, mostly on the vague basis of typology (see Greene 1966, Eska 2010b, Broderick 2013). Regrettably, too little is known about the early history of the non-IE languages with which Celtic-speakers were in contact, for example, the language of the non-Celtic Pictish inscriptions in Scotland (Jackson 1980) or Aquitanian/Basque in France and Spain. (2) In reacting to this trend, Dillon, Watkins, Meid and others stressed the degree to which Celtic could be explained by "forward reconstruction" from IE (Watkins 1962: 7); but at the same time they regarded Celtic as particularly archaic, preserving hypothetical IE features lost in other less "marginal" IE dialects (e.g. the original distribution of primary and secondary verbal endings, **Personal Endings**, p. 372). (3) Reacting in turn to this, scholars such as Rix and Cowgill argued that the peculiarities of Celtic could mostly be explained as internal developments and did not require significant modification of the traditional IE model reconstructed on the basis of Indo-Iranian and Greek (see McCone 1986: 222–223). (4) Other linguists still insisted "on the high importance of Old Irish and other genetically western branches for the reconstruction of proto-Indo-European" (Hamp 1987). Probably there is truth in all these approaches.

Celtic cannot be grouped with any other branch in an IE subdialect, unless arguably with Italic. The long-held theory of “Italo-Celtic” has had periods out of favour (see Watkins 1966, Schmidt 1991) but remains resilient (see Eska 2010a: 22, Weiss 2012, Schrijver 2015). Some of the similarities, such as the retention of the mediopassive in *-r* (shared with Hittite and Tocharian), can be regarded as shared archaisms. Others, such as the apparent equivalence of the Latin future in *-bo* and the Old Irish *f*-future, may be illusory (**Future Stem**, p. 375). Others, such as the *o*-stem genitive singular in *-ī* (also in Messapic, but not found in Celtiberian (p. 360) or Osco-Umbrian), and superlatives of the type Osc. *nessimas* ‘proximae’: Gaul. *neddamon* (gen. pl.), OIr. *nessam*, OBret. *nesham* ‘nearest’, may be due to diffusion over the long period during which Italic and Celtic speakers lived side by side (cf. Eska 2010b: 546). (In fact, the Italic influence on Celtic has never ceased, owing to continual borrowings from Latin, which have had a fundamental influence on Celtic word formation, as well as supplying many loan words.) The shortage of striking Italo-Celtic phonological developments (e.g. Watkins 1966, Zair 2012: 271) has been seen as a strong argument against the Italo-Celtic theory.

There is much less evidence for any deep connection between early Celtic and Germanic, despite the fact that they were spoken in close proximity by the time of Julius Caesar and Tacitus. There are no shared innovations in phonology and morphology, and the number of significant lexical correspondences was formerly exaggerated (see Campanile 1970, Evans 1981, Polomé 1983, Schmidt 1984, 1986b: 205–206, Schumacher 2009). According to the minimal approach of Evans (1981: 248), only about a dozen early Germanic words are certainly borrowed from Celtic, for example, Goth. *reiks* ‘ruler’: Gaul. *-rīx* < IE \*rēǵs (cf. Chapter 1); Goth. *eisarn* ‘iron’: Gaul. *Isarnus* (personal name), OIr. *iarnn*, W *haearn* (< \*(h)ǝi(h)arn-) < \*isarno-; OHG *ledar* ‘leather’: OIr. *lethar*, W *lledr* < Celt. \*l(e)tro- < IE \*pl(e)-tro- (cf. Lat. *pellis*, etc.). The Insular Celtic languages were, of course, influenced by Germanic through English at a later stage, and there is also a slight Old Norse element.

The fact is that Celtic shares *some* isoglosses with almost all other IE languages, and these can be selected to support many different theories (cf. Schmidt 1985, 1986b: 202–206, Stalmaszczyk & Witczak 1995). Similarly, scholars wishing to stress the peripheral and/or archaic nature of Celtic can and have noticed certain remarkable retentions, for example, masculine and feminine forms of the numerals ‘three’ and ‘four’ in Modern Welsh: *tri chi* ‘three dogs’, *tair cath* ‘three cats’; *pedwar ci* ‘four dogs’, *pedair cath* ‘four cats’. Generalizations about the affiliations of Celtic and about its archaic or innovative tendencies are usually subjective, lacking a statistical basis. Statistical attempts to compare the IE languages on the basis of “core vocabulary” (Elsie 1990: 318) or of occurrences under Pokorny’s roots (Bird 1982: 119–120) are open to objections, but do agree on Celtic’s closeness to Germanic and (to a lesser extent) to Latin/Italic, and a phonological link between the three, attributed to some version of Dybo’s Law, has been seen in the pre-tonic vowel shortening exemplified by OIr. *fer*, OEng. *wer*, Lat. *uir* ‘man’ versus Skr. *vīrá-* (e.g. Matasović 2012). Some form of the hypothesis of “Late Western Indo-European” (Meid 1968: 53) could explain such convergence. A more selective and rigorous comparison favours an especial closeness to Italic (see Eska 2010a: 22).

### Subdivision

The internal subdivision of the Celtic languages has not yet been agreed by scholars; either the data are inadequate, or the reality is more complex than a simple genetic model would allow.

A traditional division distinguishes “P Celtic”, in which IE \*k<sup>w</sup> > \*p, and “Q Celtic”, in which \*k<sup>w</sup> remained: Gaelic and Celtiberian are Q Celtic, whereas Brythonic, Lepontic

and most (but perhaps not all) Gaulish material is P Celtic. In view of the originally allophonic nature of the *\*k<sup>w</sup>/\*p* alternation in Celtic (**Consonantism**, p. 363), most linguists would now agree that “the ‘isogloss’ between P-Celtic and Q-Celtic is structurally trivial” (Watkins 1966: 32 n. 7, following Hamp 1958: 211). Nevertheless, the P/Q division has had a great influence on archaeological speculation, and “P” and “Q” remain useful labels for the important and valid distinction between Brythonic (or possibly “Gallo-Brythonic”) and Gaelic.

Some linguists use the term “Insular Celtic” (see **Introduction**, p. 352) in a more than geographical sense, emphasizing a special relationship between Gaelic and Brythonic (Greene 1966, McCone 1986: 262; cf. Evans 1988: 219). It is true that these groups share certain important developments – such as the development of morphophonemic mutations of initial consonants (**Consonantism**, p. 364), the system of absolute and conjunct verbal endings (**Personal Endings**, p. 372) and periphrastic constructions with preposition + verbal noun such as OIr. *oc precept*, *W yn pregethu* ‘preaching’ (Sims-Williams 2015) – but it could be argued that these developments might also have occurred in Gaulish if this were attested after AD 500 (Sims-Williams 1984: 147–148). Logically, it is impossible to establish a genetic relationship between two dialects on the basis of a shared innovation occurring after all other dialects have died out. An opposing approach, which still leads to the same negative conclusion, is to argue that a feature like the absolute/conjunct system occurs in Insular Celtic but not (apparently) in Continental Celtic because the Continental dialects were less archaic, having drifted towards Greek and Latin, or even Late IE innovations (Meid 1986: 120–121). The fundamental problem, however, is the chronological disjuncture between our evidence of Insular and Continental Celtic. A single example illustrates this. We noted above (**Introduction**, p. 353) that *\*wer* (< *\*uper*) became *\*wor* in Brythonic and Gaelic, whereas *Ver-* is attested in Gaulish. Is this innovation an Insular Celtic isogloss (Schrijver 1995: 129, 464)? The change *\*wer* > *\*wor* (due to influence of *\*wo* < *\*upo* ‘under’ or simply to rounding after /w/, cf. Evans 1967: 279) must in fact be late in Brythonic, since in AD 725 Bede, *Chronica maiora* 434, records the name of a fifth-century Briton as *Uertigernus* (> later *Uortigern*, *Gwrtheyrn*). As the change *\*wer* > *\*wor* did not occur before the fifth century AD in Brythonic, it is possible that it would also have occurred in fifth-century Gaulish as well, if this was attested – for there is evidence of shared phonetic developments between British and Late Gaulish (Fleuriot 1978). Moreover, /wer/ > /wor/ might occur quite independently; for example, there is evidence for it in Spain (Tovar 1986: 89).

The term “Gallo-Brythonic” emphasizes the similarities in language (e.g. name formations) between Gaulish and British, which were already noted by ancient writers (**Introduction**, p. 353). It is debatable whether these similarities are due to the many ethnic movements across the English Channel (already mentioned by Julius Caesar) or go back to a genuine “Gallo-Brythonic” dialect of Celtic, genetically distinct from Gaelic, Celtiberian, Lepontic and so on (cf. Evans 1988: 220, Fleuriot 1988, Lambert 1994: 17–19, Schrijver 1995: 463–465). The evidence for both the “Gallo-Brythonic” and the “Insular Celtic” hypotheses can be undermined by postulating “areal” developments (Sims-Williams 2007b: 1–42, Stifter 2008, Eska 2010a: 24–25, Lambert 2010).

An ambitious attempt at a Celtic family tree by Schmidt (1988: 235), combining the ideas of “P Celtic” and “Gallo-Brythonic” with a distinction between an “*em/en-* language” (i.e. Gaelic), which diverged very early, and the “*am/an-* languages” (i.e. Celtiberian and most or all the P Celtic languages), failed to convince (cf. Tovar 1986: 84 n. 3, Evans 1983: 29–31, McCone 1991: 22, 48–50, 161). It was based on the treatment of IE /m/ and /n/ in initial position and before plosives, e.g. IE *\*m̥bʰi* ‘about’ > Clb. *amPi-*,

Gaul., *ambi-*, *W am* versus OIr. *imb* (see further de Bernardo Stempel 1987: 38, 51, 121). A secondary Gaelic development of /-an/ > /-en/ was suggested by Cowgill (1975: 49), and some scholars postulated a Proto-Celtic allophone [æ] before nasals (e.g. Hamp 1965: 225, Joseph 1990: 126 n. 10). McCone (1996: 67–79) argued convincingly that /m/ and /n/ gave CC /am, an/ and that raising of the /a/ to /e/ or /i/ occurred in Primitive Irish (cf. Schrijver 1993).

The earliest connected written material in Celtic languages comes, as would be expected, from the Mediterranean world – Italy, France and Spain – in the second half of the first millennium BC (slightly earlier in the case of some northern Italian inscriptions: Uhlich 2007: 379). For other areas at this period there are only place-, personal and ethnic names (e.g. Sims-Williams 2007a, 2012b), but even these are of some value. For example, judging by the onomastic evidence, the language of the Galatians in Asia Minor was similar to Gaulish (Eska 2013a). Again, there is obvious Gaulish material in the “Thracian” and other eastern onomastic corpora (Falileyev 2009, 2013, 2014). Here, however, I shall concentrate on the Celtic languages from which more than names and odd words are known (see also Eska & Evans 2010).

*Lepontic* is the name given to the language of inscriptions first found within a 50 km radius of Lugano, in north-western Italy and Switzerland, and arguably distinct from Cisalpine Gaulish (Lejeune 1970, Karl & Stifter 2007: 45–73, Uhlich 2007, Stifter 2008: 285–286, Eska 2010a: 24, Eska & Evans 2010: 35). They date from the sixth to the first century BC and are written in the Lugano alphabet. This alphabet does not distinguish between /p t k/ and /b d g/ and avoids double letters; for example, the Lepontic personal name *ANOKOPOKIOS* corresponds to Gaul. *Andocombogios*. Here /nd/ > /nn/ is a Lepontic peculiarity, but the suppression of the nasal in *KO(m)P* may be purely graphic. Like Gaulish (**Consonantism**, p. 364), Lepontic distinguishes between two sibilants, both seen in *ÍSOS* [?'i'sos] < \*istos ‘that (man)’ in an inscription from Vergiate, now in the archaeological museum in Milan (Lejeune 1970: 444–452, Eska & Mercado 2005): *PELKUI PRUIAM TEU KARITE ÍSOS KALITE PALAM* = ? *Belgūi bruyyam Dēwū garite, isos kalite palam* = ? ‘Dēwū enclosed the construction for Belgos; he erected the stone.’ The word for ‘construction’, the accusative of \*bruywā, recalls Gaul. *brīvā* ‘bridge’ < \*b<sup>h</sup>rēwā, but the closest cognates are in Germanic, for example, OS *bruggia* ‘bridge’ < \*b<sup>h</sup>ruw-yo-. The word possibly meaning ‘stone’, *palā*, which is common in Lepontic inscriptions, is of unknown etymology. The verbal stem \*gar- may derive < \*gr̥-: IE \*g<sup>h</sup>er- ‘to enclose’, as in OIr. *gort* ‘field’, *W garth* ‘enclosure’ (but cf. Hamp 1991); and \*kal- may derive < \*k̥l̥-: IE \*kelH- ‘to raise, rise’, cf. Gaul. *celicon* ‘building’, Lat. *collis*, OEng. *hyll* ‘hill’ and possibly *Celtī* ‘?exalted ones’. The dental preterites are obscure, but seem to be confirmed by the Gaulish verbs *KarniTu* (3 sg.), *KarniTus* (3 pl.) found farther south in Gaulish inscriptions in Lugano script at Todi and at San Bernadino di Briona (Eska 2007/8). In the Todi inscription, which is bilingual, *KarniTu* corresponds to *LOCAVIT ET STATVIT*.

*Celtiberian* is the only Hispano-Celtic language known from inscriptions as well as proper names. (The Celticity of the Lusitanian and “Tartessian” inscriptions is generally rejected.) The “Celtiberian” inscriptions come from north-eastern Spain, in exactly the same region that the *Celtiberi* are located by ancient writers (de Hoz 2007, 2010–). The earlier inscriptions, such as the most famous one from Botorrita 20 km south of Zaragoza (c. 100 BC), are written in the native Iberian script (Eska 1989, Eichner 1989: 23–55 (= Karl & Stifter 2007: 9–44), Meid 1994a, Prósper 2008, de Bernardo Stempel 2010). This did not distinguish voiced and voiceless stops, like the Lugano alphabet, and it had the added ambiguity of using syllabic characters for these sounds (*Pa, Ca, Ta, Pe, Ce, Te*, etc.); e.g. the *Ti* symbol could denote /ti/, /di/, /t/ or /d/. Its *ś* = /s/, but

*s* = /z/ or /ð/ by lenition of \*s and \*d according to Villar (1995); for further theories see Stifter 2008: 271–273. Inevitably, then, interpretations of Celtiberian writings are disputed. Later inscriptions are in the Roman alphabet, which should be less ambiguous; however, opinions still diverge. For example, Ködderitzsch (1985) offered the following reading and interpretation of the first- or second-century AD inscription from Peñalba de Villastar (cf. Villar 1991): *ENIOROSEI VTA TIGINO TIATVNEI [-MEI?] ERECAIAS TO LVGVEI ARAIANOM COMEIMV ENIOROSEI EQVEISVIQVE OGRIS OLOGAS TOGIAS SISTAT LVGVEI TIASO TOGIAS* ‘To Enior(o)sis and to Tiatū of Tiginos we bestow furrows, and to Lugus a field; to Enior(o)sis and to Equaesos, Ogris submits the protections of the fertile-land, to Lugus the protections of the scorched-land.’ Meid (1994a and 1994b) and others, however, offer completely different interpretations (see Jordán Coléra 2004: 375–390; de Bernardo Stempel 2008). Thus, Meid (1994a: 36–37) translates: ‘To the mountain-dweller as well as to . . ., to Lugus of the Araians, we have gone on a field-procession (reading *TRECAIAS*); for the mountain-dweller and horse-god, for Lugus, the head of the community has put up a covering (i.e., a house or a hall), (also) for the *thiasus* a covering’. And Prósper (2002) has ‘In Orosis and as far as the Tiginos reaches, to Lugu we consecrate the fields. In Orosis and in Equiso, the hills as well as the ploughed fields and the houses are consecrated to Lugu, that is, the houses of the bounded area’. Notable phonological points here include:

1. the loss of /p/ (if Ködderitzsch’s etymologies are correct) in (a) *er(e)caia-* ‘furrow’ < IE \*perk- (cf. W *rhych*, Lat. *porca*, Eng. *furrow*); (b) *ol(o)ga-* ‘fertile-land’ < IE \*polǵ<sup>(h)</sup>ā, \*polkā (cf. Gallo-Latin *olca* > Fr. *ouche*, Eng. *fallow* < Gmc \*falgā); and (c) *tiaso* ‘burnt land’ < \*teposo- (cf. OIr. *tee* ‘hot’ < \*tepe-, Lat. *tepeō*);
2. the retention of IE /k<sup>w</sup>/ in *-que* ‘and’ (Lep. *-pe*, arch. OIr. *-ch*) – this is used here alongside the connective *uta*, cognate with Skr. *uta* and OBret. *ut* (Eska 1990, Sims-Williams 2007b: 234–235);
3. the apparent (orthographical?) retention of IE /ey/ in *com-(m)ei-mu*, ‘we bestow’ < IE \*mey- according to Ködderitzsch (cf. Skr. *máyate* ‘exchanges’, Lat. *mīnus* < \*moy-nes-, OIr. *moín* ‘treasure’ < \*moy-ni-, MW *mwyn* ‘value’ < \*mey-no-) or < IE \*h<sub>1</sub>ey- ‘go’ according to Meid, Prósper and de Bernardo Stempel;
4. the retention of final /m/ (not > /n/) in the acc. sg. *ar(a)ianom* (‘field’, according to Ködderitzsch, cf. OIr. *airim*, MW *ardaf* ‘I plough’ < \*ary-, or, according to de Bernardo Stempel, a neuter adjective related to W *iawn* ‘just’) – Meid’s gen. pl. ethnonym in *-om* (not *-um!*) is unlikely;
5. in *sistat* ‘puts’ < \*sistati (or 3 pl. \*sistanti according to Prósper) the apocope of primary \*-i, as in Lat. *sistit* (**Personal Endings**, p. 372–373) – although some analyse *sistat* as if = \*sistaz < old secondary \*sistat;
6. epenthesis in *er(e)caias*, *ol(o)gas*, *ar(a)ianom*, and *Enior(o)sei* (Eska 1996).

In morphology note the *o*-stem genitives in *-o* (*Tigin-o*, *tias-o*) rather than the *-ī* found elsewhere in Celtic and Italic (**Position**, p. 357; cf. McCone 1994: 96 = 2006: 83, Eska 1995, Schmidt 1995: 252–253).

*Gaulish* inscriptions begin in about the third century BC, but the bulk of material comes from the first century BC to the second century AD (Stifter 2012). The earlier material is mostly written in Greek script, with a few texts in the Lugano script from Cisalpine Gaul. The later material is mostly in Roman letters. To the inscriptions on stone which have long been known can now be added important long texts on metal plates, notably from Chamalières, Larzac and Chartres (Lambert 2002, Viret *et al.* 2013); these are

transforming, and confusing, our understanding of the Gaulish language(s). The following example of Gaulish is a stone inscription from Alise-Sainte-Reine (first century AD): *MARTIALIS DANNOTALI IEVRV VCVETE SOSIN CELICNON ETIC GOBEDBI DVGHIONTHIO VCVETIN IN ALISIA* (Lambert 1994: 98–101, Eska 2003) ‘Martialis [son] of Dannotalos offered to [the god] Ucuētis this edifice, and to the smiths who honour (?) Ucuētis in Alisia’. Note here *-e* < *\*-ey* in the dat. *Ucuete* and *-n* < *\*-m* in the acc. sg. *celicnon* (the source of Goth. *kelikn* ‘tower’). Morphological points of interest include the genitive singular in *-i* in *Dannotali*, the dative (or instrumental?) plural in *-bi* in *gobedbi*, the third-person singular preterite in *-u* (cf. Lepontic) and the indeclinable relative particle *io* in *dugiionti-io* (**Pronouns**, p. 370).

Apart from two curse-tablets from Roman Bath (Mullen 2007), *British* and its successors, Primitive Welsh, Primitive Cornish and Primitive Breton, are known only from proper names in inscriptions and Latin texts, which provide, nevertheless, detailed information about phonological developments (Jackson 1953, Sims-Williams 1990, 1991, 2003). From c. AD 800 onwards we have manuscript glosses, memoranda and so on in Old Welsh, Old Cornish and Old Breton (Ternes 2011). (In addition, some Welsh poetry may be as early as the sixth century, although transmitted in late manuscripts.) The orthography of Old Welsh, Old Cornish and Old Breton (OWCB) is based on a pronunciation of Latin found in Britain, the main feature of which was that medially Latin consonants underwent “lenition” (**Consonantism**, p. 364), so that Latin words like *medicus*, *decimatus* were pronounced [meðigəh], [degiµa:dəh]. Consequently, similar values were assigned to letters in writing OWCB; for example, [degued] (‘tenth’ < *\*dekametos* (see Table 7.1 on p. 353)) would be written *decmet*.

*Primitive Irish* (McCone 1994, 1996, 2005) is known from proper names in inscriptions and Latin texts, and its phonological development is clear in outline (McCone 1996, Sims-Williams 2003: 296–301). Most Irish inscriptions of the fifth, sixth and seventh centuries AD are in the Ogham alphabet (McManus 1991, Ziegler 1994, Sims-Williams 2007b). This is well suited to carving on wood and stone, since only straight strokes are used, e.g. *////* = /k<sup>w</sup>/, *////* = /k/. The phonology of the earliest Ogham inscriptions is archaic, for example, distinguishing between /k/ and /k<sup>w</sup>/ and (probably) between /g/ and /g<sup>w</sup>/, and showing the old case endings. The earliest extant manuscripts containing *Old Irish* (glosses and short texts) are eighth century, but it is probable that some texts extant in late manuscripts were written down towards the end of the sixth century. Most Old Irish material is written in an orthography based on the British pronunciation of Latin described above (for exceptions see Harvey 1989, Russell 1995: 224); hence, medial and final [b d g] are written *p t c*, and medial and final [β δ γ µ] are written *b d g m*. Medially and finally double consonants may be used to avoid ambiguity; e.g. [b] may be written *bb*, and [k] may be written *cc*. Palatalized consonants are indicated by flanking vowels, e.g. *macc* [mak], *maicc* [makʲ], *beirid* [bʲerʲəðʲ], *feraib* [fʲerəβʲ]. In modern edited texts, diphthongs are generally distinguished by placing the length mark on the *i*; e.g. *ai* and *úi* are diphthongs (the former also written *áe/oí/óe*), but *ái* and *úi* denote [a:] and [u:] followed by palatalized consonants (in the phonetic transcription <sup>ʲ</sup> marks a palatalized consonant, as in OIr. *nemed* [nʲeµʲəð]), **Introduction**, p. 354).

## THE PHONOLOGY OF COMMON CELTIC

The following features would generally be accepted as defining “Celtic”. For ease in consulting handbooks (e.g. Lewis & Pedersen 1961), I give IE sounds in their traditional reconstructions without prejudice to the actual phonetic reality of /b<sup>h</sup>/, etc. Also, not much

attention is paid to laryngeals, for which Celtic provides little independent evidence (cf. Hamp 1965, Joseph 1982, Ringe 1988, Lindeman 1988, 1989: 291–293, McCone 1994: 71–73 = 2005: 35–38, Zair 2012). Almost nothing certain is known about the fate of the IE free accent in Celtic nor about the development of the very different initial and penultimate stress accents of Gaelic and Brythonic respectively (cf. Koch 1987, Lindeman 1992, Schrijver 1995: 16–22). There is already evidence for penultimate accentuation in Continental Celtic (de Bernardo Stempel 1995).

### Vocalism

The IE short vowels /i e a o u/ remained in Common Celtic. A laryngeal regularly fell together with /a/, e.g. Gaul. *atir* (Larzac inscription), OIr. *athair* ‘father’ [aθ̪əɾʲ]. The IE long vowels /ī: a: u:/ remained in Common Celtic, but IE /e:/ > /i:/, e.g. \*r̥ēgs (: Lat. *rēx*) > Gaul. *-rix*, OIr. *rí*, W *rhi*, and IE /o:/ > /a:/, e.g. \*d̥ōnom (: Lat. *dōnum*) > OIr. *dán* ‘gift’, W *dawn* (with /au/ < /ɔ:/ < /a:/) ‘gift’, except when IE /o:/ occurred in final syllables, where it became CC /u:/, e.g. \*b<sup>h</sup>erō > \*birū > OIr. *-biur* ‘I carry’ and Gaul. *delgu* ‘I hold’ (cf. Eska 1995: 34). One gap in the resulting Common Celtic system of vowels (shown in Table 7.2).

TABLE 7.2 THE COMMON CELTIC VOWEL SYSTEM

i: i		u u:
e		o
	a	
	a:	

was filled by the development /ei/ > /e:/ (cf. Gr. *steikhein* ‘to walk’: Early OIr. *-tēgot* [t̪e:γod] ‘they go’ = Later OIr. *-tiagat* [t̪iayəd]), but /ei/ > /e:/ may not have been completed in Common Celtic. (See p. 362 for Clb. *ei*, and note dat. *-ei* in Lepontic as well.) The other gap was filled in Gaelic by /eu au ou/ > /o:/ (later alternating with /uə/) and in Brythonic by /eu ou/ > /o:/ (later > /u:/), but in Common Celtic /eu ou au/ remained diphthongs, as did /ai/ and /oi/. There was, however, an early tendency, seen already in Lepontic and Gaulish, for /eu/ > /ou/, e.g. \*teutā ‘people’ > *touta* (> OIr. *túath* [tuəθ], MW *tut* [t̪u:ð]).

The semi-vowels /w/ and /y/ remained in Common Celtic, and indeed survive to this day in Welsh. Vocalic /m̥, n̥, r̥, l̥/ developed as vowel + consonant (/am, em, an, en, ar, al/) or consonant + vowel (/ri, li/), depending on context (see de Bernardo Stempel 1987, McCone 1996: 67–79); there is often apparent divergence between /em en/ in Gaelic and /am, an/ elsewhere (**Subdivision**, p. 358). The so-called long resonants, that is, vocalic /m̥, n̥, r̥, l̥/ + laryngeal /H/, developed mostly as /ma:, na:, ra:, la:/, e.g. IE \*ǵrH-no- > OIr. *grán*, W *grawn* ‘grain’ (: Lat. *grānum*). The derivation of reflexes with short /a/ is debated: for example, does OIr. *tarathar*, W *taradr* ‘auger’ < \*tara-tro-n come from \*t̪rH-, or is *tara-* due to vowel harmony in \*tera-tro-n coming from \*terH-? The latter is probable. (See Joseph 1982, de Bernardo Stempel 1987: 43–45, Lindeman 1988, Schrijver 1995: 87.)

### Consonantism

/m, n, r, l, s/ remained basically unchanged (on *s* see Stifter 2012). Final *-m* became *-n* in most Celtic languages, but not in Lepontic, Celtiberian (p. 359–360) and some

Gaulish. As in other IE languages, /s/ had an allophone [z]; this is not differentiated from *S* in Continental Celtic writings, but in Insular Celtic [z] > [δ], e.g. Gaul. *TASGO*:- OIr. *Tadg* (personal name). As in Armenian, IE /p/ was completely lost, apparently via /f/ (Eska 2013b); it had already been reduced to /h/ or /zero/ when classical writers borrowed the name of the *Hercynia silva* in central Germany (: IE \*perk<sup>w</sup>us ‘oak’, cf. Evans 1979: 531–532). Another commonly cited example is the classical name for Ireland/the Irish, *Ierne/Hierni*, but it is doubtful whether this name is Celtic and cognate with Skr. *pīvarī*, Gr. *πίερα* ‘fat, rich’ (Schrijver 2014: 77–78). Among the plosives there is no trace of an original palatal series (Gaelic palatalization arose much later, p. 367), and the only trace of the aspirate/unaspirate distinction is between the labiovelars /g<sup>w</sup>/ > CC /b/ and /g<sup>wh</sup>/ > CC /g<sup>w</sup>/ (Cowgill 1980, Sims-Williams 1981, 1995). For IE /g<sup>w</sup>/ note OIr. *béo*, W *byw* ‘living’: Lat. *uīuus*; OIr. *imb* ‘butter’: Lat. *unguen*; the only certain exception is before /y/, which delabialized /g<sup>w</sup>/, e.g. OIr. *nigid* [n<sup>i</sup>ɣ<sup>i</sup>əδ<sup>i</sup>] ‘washes’ < \*nig<sup>w</sup>-ye-ti and W *gīau* ‘sinews’: Ved. *j(i)yā-* ‘bowstring’, Av. *jīiā-* ‘bowstring, sinew’, Gr. *βίος* ‘bow’. For IE /g<sup>wh</sup>/ note OIr. *gonaid* ‘wounds’, W *gwanaf* ‘I wound’: Gr. *φόνος* ‘murder’, *θείνω* ‘I strike’, Hitt. *kuenzi*, Skr. *hanti* ‘kills’. By the time of Old Irish, /g<sup>w</sup>/ was delabialized, but there is probably an old symbol for /g<sup>w</sup>/ in the Ogham alphabet (p. 361). In Brythonic, initial /g<sup>w</sup>-/ > /gw-/, probably directly and not via /w-/ in view of the early Welsh inscription *GVANI*, which predates the general change of IE /w-/ > W /gw-/ (Sims-Williams 1995). In Gaulish /g<sup>w</sup>-/ > /w-/ is possible, if the first-person singular verb *uediiumi* (? *uediiu mi*) in the Chamalières inscription is cognate with W *gweddi* ‘prayer’, OIr. *guide* ‘prayer’ < ?\*g<sup>w</sup>edyā, Gr. *ποθέω* ‘I wish’, Av. *jaidiia-* etc. (Cowgill 1980). The fate of non-initial /g<sup>w</sup>/ < /g<sup>wh</sup>/ is debated. In Welsh it appears to have had the same result intervocally as /g/, i.e. loss (e.g. MW *de* ‘burns’ < IE \*d<sup>h</sup>eg<sup>wh</sup>-). The words *nyf* ‘snow’ and *deifio* ‘to burn’ (*f* = [v]) are probably wrongly cited in the handbooks as reflexes of IE intervocalic /g<sup>wh</sup>/; *nyf* probably shows a special treatment of g<sup>wh</sup> after /n/ (cf. Lat. *ninguis*), while *deifio* is not from IE \*d<sup>h</sup>eg<sup>wh</sup>- at all (Sims-Williams 1995).

The above data are best explained by the following chronology (cf. Sims-Williams 1981: 227), starting from the system

TABLE 7.3 PRE-CELTIC STOPS

(p)	t	k	k <sup>w</sup>
b	d	g	g <sup>w</sup>
b <sup>h</sup>	d <sup>h</sup>	g <sup>h</sup>	g <sup>wh</sup>

Common Celtic merged /g<sup>w</sup>/ and /b/ as /b/. Then de-aspiration, occurring throughout the system (/b<sup>h</sup>/ > /b/, /d<sup>h</sup>/ > /d/, etc.), created a new /g<sup>w</sup><sub>2</sub>/ < /g<sup>wh</sup>/ . The relative chronology of /p/ > /zero/ is uncertain, but it was the gap in the system resulting from the loss of /p<sub>1</sub>/, which made possible the allophonic alternation between “Q Celtic” with [k<sup>w</sup>] and “P Celtic” with [p<sub>2</sub>] < [k<sup>w</sup>] (Subdivision, p. 357):

TABLE 7.4 COMMON CELTIC STOPS

[p <sub>2</sub> ]	t	k	[k <sup>w</sup> ]
b	d	g	g <sup>w</sup> <sub>2</sub>



The most interesting combinatory change among the consonants was the development of a new dental phoneme, written  $\delta\delta$ ,  $ss$  and so on, in Gaulish (and apparently with special symbols in Lepontic, p. 359), from /d/ + /t/, /t/ + /t/, /t/+ /s/, etc. (Evans 1967: 410–420). Note also a general development of /xt/ < /pt, kt/. Structurally, a more important development was the widespread rise – as in Germanic – of intervocalic geminate consonants contrasting with single consonants in the same position (Kuryłowicz 1960: 259–273, de Bernardo Stempel 1989). Such geminates are often shown in Gaulish, but are excluded by the scripts used for Lepontic and Celtiberian. In Insular Celtic geminates developed differently from single consonants, for example, OIr. *maicc* [makʲ] ‘of a son’ < Ogham *MAQQI* = \*mak<sup>w</sup>k<sup>w</sup>ī, but MW *meib* ‘sons’ < \*mapī < \*mak<sup>w</sup>ī; the single /p/ is “lenited” (i.e. voiced), but the geminate /k<sup>w</sup>k<sup>w</sup>/ is simplified. In addition to the *phonemic* distinction between /VCV/ and /VCCV/, there arose, probably already in Common Celtic, an *allophonic* distribution of [C] and [CC] in other environments (Harvey 1984). The evidence for this comes partly from the phonemicization of reflexes of these allophones in Insular Celtic (**Consonantism**, p. 365, 367), partly from the analogy of Romance, which may reflect Celtic substratum influence here (Martinet 1952; cf. Villar 1995). Ultimately, this was of great importance in external sandhi, giving rise to the Insular Celtic system of *initial mutations* (Russell 1995: 231–257, Sims-Williams 2007b: 43–58). For example, \*esyō kattos ‘his cat’ > OIr. *a chatt* [ə xat] (lenition), MW *y gath* [i ga:p] (lenition), but \*esyās kkatus ‘her battle’ > OIr. *a cath* [ə kaβ] (no mutation), MW *y chat* [i xa:d] (spirant mutation). A further set of mutations was produced after old nasals (**Consonantism**, p. 366), e.g. OIr. *a catt* [ə gat] ‘their cat’, MW *vyg cath* [və ŋha:p] ‘my cat’. These initial mutations, which began as sandhi phenomena, were grammaticalized in Insular Celtic, for example, as markers of relative clauses (Ó hUiginn 1986).

## THE PHONOLOGY OF EARLY BRYTHONIC

The accent in later British fell on the penultimate syllable, which became the ultimate syllable after the loss of final syllables *c.* 500. A full range of vowels was therefore preserved in the final syllables of OWCB words, whereas pre-tonic vowels tended to have been shortened, reduced or syncopated. (Much later, by *c.* 900, at least in Welsh (Sims-Williams 2003: 289), the accent shifted from the ultimate to its present position on the penult, except in the Vannetais dialect of Breton.) Schrijver 1995 and Ternes 2011 are the most up-to-date handbooks.

### Vocalism

IE vowel quantity was at first retained, but by the sixth century AD a new quantity system applied automatically even in stressed syllables; according to this, Primitive Welsh, Cornish and Breton vowels were short in [VCC] syllables and long in [V(C)] syllables (Sims-Williams 1990: 250–260). In *stressed* syllables, the Common Celtic vowels and diphthongs developed mainly as follows:

/i/ > Pr. W /i(:)/ (written <y> in later W), Pr. Corn., Pr. Bret. /i(:)/

/e/ > Pr. W/Corn./Bret. /e(:)/

/a/ > Pr. W/Corn./Bret. /a(:)/

/o/ > Pr. W/Corn./Bret. /o(:)/

/u/ > Pr. W/Corn./Bret. /u(:)/, also Pr. Corn./Bret. /o(:)/ (note: /u(:)/ is written <w> in Welsh, <ou> in Breton)

- /i:/ (< IE /e:/ and /i:/) > Pr. W/Corn./Bret. /i(:)/
- /e:/ (< CC /ei/) > Pr. W/Corn./Bret. /ui/
- /a:/ (< IE /a:/ and /o:/) > British /ɔ:/ > Pr. W /au/, Pr. Corn./Bret. /ə(:)/
- /u:/ (< IE /u:/ and /-o:/) > British /u:/ > Pr. W/Corn./Bret. /i(:)/
- /au/ > British /ɔ:/ > Pr. W /au/, Pr. Corn./Bret. /ə(:)/ (see Schrijver 1995: 195)
- /ou/ (< CC /ou/ and /eu/) > British /o:/ > /u:/ > Pr. W/Corn./Bret. /u(:)/
- /ai/ > British /e:/ > Pr. W/Corn./Bret. /oi/
- /oi/ > British /u:/ > Pr. W/Corn./Bret. /u(:)/

The semi-vowel /w/ had become /gw/ in absolute anlaut by the time of the earliest OWCB (c. 800). Medial /y/ sometimes developed to /ð/.

**Consonantism**

/s/ tended to become /h/ or /y/ or to disappear. [z] became /ð/, and [x] was vocalized as /y/. (On /g<sup>w</sup>/ see p. 363 above.) Most other consonants suffered the change known as “lenition” in positions where their weaker allophones occurred (cf. p. 364), and there was later a tendency for unlenited voiceless consonants to be spirantized (except in absolute anlaut):

**TABLE 7.5 BRYTHONIC CONSONANTAL DEVELOPMENTS**

[pp]	>	[p]	> [f]	(Spirantization)
[p]	>	[b]	(Lenition)	
[tt]	>	[t]	> [β]	(Spirantization)
[t]	>	[d]	(Lenition)	
[kk]	>	[k]	> [x]	(Spirantization)
[k]	>	[g]	(Lenition)	
[bb]	>	[b]		
[b]	>	[β]	(Lenition)	([β] later > [v])
[dd]	>	[d]		
[d]	>	[ð]	(Lenition)	
[gg]	>	[g]		
[g]	>	[γ]	(Lenition)	([γ] later lost, or > [y] or [w])
[mm]	>	[m]		
[m]	>	[μ]	(Lenition)	([μ] later > [v])

These changes probably took place in three stages: (a) spirantization of /b, d, g, g<sup>w</sup>, m/ (before c. AD 400?); (b) voicing of /p, t, k/ (fifth century?); and (c) spirantization of /p<sub>2</sub>, t<sub>2</sub>, k<sub>2</sub>/ (sixth century?) (see Sims-Williams 2007b: 14–16, 26, 43–58). Since they occurred in external sandhi, they led to “lenition” (e.g. /k ~ g/) and “spirantization” (e.g. /k ~ x/) as initial mutations (p. 364). The above simplifications of double consonants (except [mm], which followed the pattern of [nn]) were completed before the advent of the new quantity system (**Vocalism**, p. 364), hence, for example, W *crēd* ‘belief’ < \*krēdd- (versus *mām(m)* ‘mother’ < \*māmm-).

**THE PHONOLOGY OF GAELIC**

The accent in Gaelic fell on the initial syllable of stressed words. A full range of vowels was therefore preserved in this syllable, whereas post-tonic vowels tended to be shortened, reduced or syncopated. For example, in Old Irish there were only two short vowels

in closed unstressed syllables, [ə] and [u] (although the spelling system seems to disguise this). Because Brythonic stress developed very differently (p. 364), the evidence of the two branches is of complementary importance in reconstruction. The best handbook is McCone 1996; cf. Sims-Williams 2003: 296–301.

### Vocalism

IE vowel quantity was retained in stressed syllables; in unstressed syllables long vowels were shortened in Primitive Irish, except in final syllables before final /h/ < /p δ x s/, and even in these syllables the vowel was eventually shortened, e.g. Celt. \*teutās ‘tribes’ > Pr. Ir. \*tōpāh > OIr. *túatha* [tuap̪ā]. In *stressed* syllables, the Common Celtic vowels and diphthongs developed mainly as follows:

- /i/ > OIr. /i/ (if not lowered to /e/ by following low vowel)
- /e/ > OIr. /e/ (if not raised to /i/ by following high vowel)
- /a/ > OIr. /a/
- /o/ > OIr. /o/ (if not raised to /u/ by following high vowel)
- /u/ > OIr. /u/ (if not lowered to /o/ by following low vowel)
- /i:/ (< IE /e:/ and /i:/) > OIr. /i:/
- /e:/ (< CC /ei/) > OIr. /e:/ alternating with /ia/
- /a:/ (< IE /a:/ and /o:/) > OIr. /a:/
- /u:/ (< IE /u:/ and /-o:/) > OIr. /u:/
- /au/ > OIr. /o:/ alternating with /ua/
- /ou/ (< CC /ou/ and /eu/) > OIr. /o:/ alternating with /ua/
- /ai/ and /oi/, though still distinct in most Ogham inscriptions, merged in Old Irish as a diphthong of uncertain value (/oi/?), written <áe>, <aí>, <óe>, <oí>.

The semi-vowel /w/ became /f/ in absolute anlaut and /v/ after nasals, for example, \*wiros > *fer* ‘man’, \*banwos > *banb* [banv] ‘pig’ (: Gaul. *Banuus*, W *banw*); in other positions /w/ was lost, as was /y/, e.g. \*yowankos (: W *ieuanc*) > \*(y)o(w)egah > OIr. *oac* [oəg] ‘young’.

### Consonantism

/s/ tended to become /h/ or to disappear. [z] became /δ/, but [x], written *ch*, remained in Old Irish. In the Ogham inscriptions /k<sup>w</sup>/ and (probably) /g<sup>w</sup>/ were still distinct from /k/ and /g/, but they had been delabialized by the Old Irish period. Even before the Ogham inscriptions the combinations /nt, nk, nk<sup>w</sup>, ns/ had become /dd, gg, g<sup>w</sup>g<sup>w</sup>, ss/, with compensatory lengthening of a preceding /a/ or /e/ to /ε:/, e.g. \*sentus (: Bret. *hent*, OHG *sind* ‘road’) > \*/sε:ddus/ > OIr. *sét* [sʲε:d] ‘road’ (cf. Ogham personal name *SEDANI* > OIr. *Sétin(a)i*); \*k<sup>w</sup>enk<sup>w</sup>e (: Lat. *quinque*) > /k<sup>w</sup>ε:g<sup>w</sup>g<sup>w</sup>e/ > /ko:gʲe/ > *cóic* [ko:gʲ] ‘five’ (the rounding was due to the labiovelars); \*Brigantī (: Skr. *bṛhatī* f. ‘exalted one’) > /brige:ddi:/ > /briγēdi/ > *Brigit* [bʲrʲiγʲədʲ] (personal name). These changes after nasals also took place in external sandhi, giving rise to the *initial nasal mutation*, e.g. gen. pl. \*wiran trumman > \*wira ddrumman > OIr. *fer tromm* [fʲer drom] ‘of heavy men’. (This mutation was not usually shown in writing.) As in Brythonic, most consonants underwent “lenition” in positions where their weaker allophones occurred (cf. p. 364), but lenition of /t/ and /k/ took a different form in Gaelic:

TABLE 7.6 GAELIC CONSONANTAL DEVELOPMENTS

[tt]	>	[t]	
[t]	>	[p]	(Lenition)
[kk]	>	[k]	
[k]	>	[x]	(Lenition)
[bb]	>	[b]	
[b]	>	[β]	(Lenition)
[dd]	>	[d]	
[d]	>	[δ]	(Lenition)
[gg]	>	[g]	
[g]	>	[ɣ]	(Lenition)
[mm]	>	[m]	
[m]	>	[μ]	(Lenition)

Gaelic lenition probably took place in two stages: (a) spirantization of /b d g<sup>(w)</sup> m/; (b) spirantization of /t k<sup>(w)</sup>/; they cannot be dated precisely, but (b) occurred later than the (fifth-century?) voicing of /p t k/ in British (see Sims-Williams 1990: 233, McCone 1994: 74 = 2005: 42). In external sandhi, lenition resulted in an initial mutation (**Consonantism**, p. 364). Internally, lenited consonants were often lost with compensatory lengthening, for example, Ogham *SAGRAGNI* (gen. sg.) > OIr. *Sárán* [sa.ra:nʲ]; Pr. Ir. \*epn- (: *Wedn* < \*petnos) > OIr. *én* ‘bird’. Note that this gave rise to new long vowels in unstressed syllables. Phonemically, the most important other change was the growth of palatalized consonants before front vowels in certain environments (Greene 1973), e.g. \*alyos (: Lat. *alius*, MW *eil*) > Pr. Ir. \*alīyah > \*alēyah > OIr. *aile* [alʲe] ‘other’. Contrast \*kaletos (: *W caled* ‘hard’) > Pr. Ir. \*kalepāh > OIr. *calad* [kaləð] ‘hard’, and note that palatalization had become phonemic at the point when [alēyah] was opposed to [kalepāh]. All consonants could be palatalized. On the spelling of palatalized consonants see *Old Irish*, p. 361.

The typical development of consonants in Insular Celtic can be exemplified by IE \*t (and the older handbooks’ \*th) as follows: IE \*t (and \*tʰ) > Celtic /t/; Celtic /t/ allophonically = [tt] and [t]; in Brythonic [t] > /t/ (absolute anlaut) and /p/ (elsewhere), but [t] > /d/; in Gaelic [tt] > /t/ and /tʲ/, but [t] > /p/ and /pʲ/. Thurneysen (1946: 97) maintained that Old Irish also had *u*-quality consonants (/tʰ/ etc.), and some favour re-instating his three-way system (Anderson 2014, McCone 2015, versus Jaskuła 2014).

THE MORPHOLOGY OF COMMON CELTIC

In nominal morphology the threefold distinctions in gender (masculine, feminine, neuter) and number (singular, plural, dual) survived from late Indo-European. The neuter was lost in Middle Irish, and only traces remain in Brythonic. The dual is always reinforced by the numeral *dá* ‘two’ in Old Irish and is merely residual in Brythonic, where it is formally identical with the singular or plural, for example, *W y gaffl* ‘the fork’ (< \*sindos gablos), *y geiffl* ‘the forks’, *Yr Eiffl* (mountain-name) ‘the (two) forks’: the last two forms imply \*sindī gablī (not du. \*\*sindō gablō), but whereas the masculine plural has dropped lenition by analogy with non-lenition after the feminine plural article *y* < \*sindās, the lenition remains in the dual. This is an illustration of “the eviction or replacement of a morph by a new morph only in the former’s primary or secondary function” (Kuryłowicz 1964: 14).

The number of cases is reduced to five in Old Irish, with a “dative” case subsuming the functions of the dative, ablative, locative and instrumental. (Some of these distinctions remain in Celtiberian and Gaulish.) Owing to syncretism, the endings of the “dative” may derive not only from the IE dative, but also from the IE ablative, locative or instrumental.

In Insular Celtic, inflected forms began to be “hypercharacterized” (Schmidt 1974), owing to the functions of the case endings being increasingly subsumed by fixed syntactic structures such as preposition + noun, VSO word order (**Syntax**, p. 377), noun + dependent genitive (in prose texts) or noun + qualifying adjective. It is not surprising, then, that case endings were often allowed to remain ambiguous in Old Irish and disappeared altogether in Brythonic, apart from a few fossils, for example, the dative of *penn* ‘head’ (*o*-stem) in MW *erbyn* ‘against, to meet’ < \*are pennū = OIr. *ar chiunn* (+ gen.) ‘id.’ < \*are k<sup>w</sup>ennū. In Brythonic, nouns (and some adjectives) have only singular and plural. Some plurals are historical, for example, MW *mab* ‘son’, pl. *meib* (form used after numerals) < \*mapos, -ī (cf. OIr. *macc*, pl. *maicc*), but many are analogical, e.g. *meib(i)on* ‘sons’, with *-(i)on* from the old *n*-stem pl. \*-ones. Ahistorical plurals were inevitable in Brythonic wherever there would have been no distinction of number; for example, \*donyos, pl. \*donyī ‘man’ (< \*gdonyos: Gr. χθόνιος, **Word Formation**, p. 377) gave *duine*, pl. *duini* in Old Irish, but *dyn*, pl. \*\*dyn → *dynion* in Welsh; plural *dyn* survived only after numerals, where it came to be regarded as singular. Many Brythonic plurals were old collectives, which may account for the use of singular verbs with plural subjects in Welsh.

The case system survived the loss of final syllables in Old Irish because the latter left traces in vowel affection, palatalization and following mutations. The paradigm of the masculine *o*-stem \*wiros ‘man’ (: Lat. *uir*) illustrates this. Parallel endings are given here (and for the *ā*-stems below) from Gaulish, Lepontic and Celtiberian (Evans 1967: 420–426, Lejeune 1970: 467, 1985a, 137–138, 1985b, Tovar 1986: 91–92, Eska 1989: 160–163, 1995, Prodocimi 1989, Lambert 1994: 49–58, Villar 1995, Eska & Evans 2010: 33, 37, 40); the quantities of vowels in these scripts are sometimes guesswork.

TABLE 7.7 NOMINAL STEMS IN -*O*-

Singular				
nom.	<i>fer</i>	[-r]	< *wiros	cf. Gaul. -os, Lep. -os, Clb. -oś
voc.	<i>fir<sup>L</sup></i>	[-r̥]	< *wire	cf. Gaul. -e(?)
acc.	<i>fer<sup>N</sup></i>	[-r]	< *wiron	cf. Gaul. -om, -on, Lep. -om, Clb. -om
gen.	<i>fir<sup>L</sup></i>	[-r̥]	< *wirī	cf. Gaul. -ī, Lep. -ī, -oiso, ?-ū, Clb. -o (see <i>Celtiberian</i> on p. 360)
dat.	<i>fiur<sup>L</sup></i>	[-r]	< *wirū	cf. Gaul. -ūi, -ū; Lep. -ūi, Clb. -ūi, -ei (loc.), -ūs (abl.)
Plural				
nom.	<i>fir<sup>L</sup></i>	[-r̥]	< *wirī	cf. Gaul. -oi, -ī, Lep. -oi, Clb. ?-oi
voc.	<i>firu</i>	[-ru]	< *wirūs	
acc.	<i>firu</i>	[-ru]	< *wirū(n)s	cf. Gaul. -o(̄)s, -ūs, Clb. -ūs(?)
gen.	<i>fer<sup>N</sup></i>	[-r]	< *wirōn	cf. Gaul. -on, Clb. -u(̄)m
dat.	<i>feraib</i>	[-rəβ̥]	< *wirobis	cf. Gaul. -obo, Lep. -oPos, Clb. -uPos

<sup>L</sup> = +lenition, <sup>N</sup> = +nasal mutation.

The dative singular *-ū(i)* may derive from the IE dative \**-ōy*, and from instrumental \**-ō* or ablative \**-ōd*. The old nominative/vocative plural \**-ōs* > \**-ūs* has been replaced by \**-ī* (< pronominal \**-oy*) in the nominative but survives in its secondary function as vocative. The Old Irish palatalization in the dative plural points to \**-bis*, an instrumental ending; cf. Gaul. *gobedbi* (p. 361). The genitive plural *fer<sup>N</sup>* points to \**-ōm*, not \*\**-ūm* < \**-ōm*; probably all long vowels before final nasal were shortened very early in Celtic, before the five long vowels were reduced to three (Cowgill 1975: 49, Jasanoff 1989: 139), although Clb. *-um* (*if* = /u:m/ < /o:m/, which is uncertain) may arguably tell against this (cf. Sims-Williams 2007b: 7–14, Eska 2010a: 23, Eska & Evans 2010: 33).

The Old Irish *ā*-stem declension is also quite well paralleled in Continental Celtic (see Table 7.8). The most controversial Old Irish ending has been the accusative singular with

palatalized *-th*, but probably the development was *\*-ām* > *\*-ǣm* (cf. above) > *\*-æm* causing palatalization (cf. **Subdivision**, p. 359). Late Gaulish borrowed *-im* from the *i*-stems. There is clear syncretism in the genitive singular, where both Irish and Late Gaulish have replaced *\*-ās* with *\*(i)yās*, the pronominal and *yā*-stem ending (cf. Lat. *pater familiās*), which developed via Ogham *-EAS* to OIr. *-e*. The original *\*-ās* was retained in the irregular paradigm of OIr. *ben* ‘woman’. This also preserved an old ablaut pattern, for example, nominative/vocative/accusative plural *mná* (= Gaul. *mnās* < *\*bnās*) and genitive plural *ban<sup>N</sup>* (< *\*banom*, cf. Gaul. *banom*, de Bernardo Stempel 1987: 83, Delamarre 2003: 73). The singular is given in Table 7.8.

**TABLE 7.8 NOMINAL STEMS IN *-Ā-***

<i>Singular</i>				
nom.	<i>túath<sup>L</sup></i>	< <i>*teutā</i>		cf. Gaul. <i>-ā</i> , Lep. <i>-ā</i> , Clb. <i>-ā</i>
voc.	<i>túath<sup>L</sup></i>	< <i>*teutā</i>		cf. Gaul. <i>-ā</i>
acc.	<i>túath<sup>N</sup></i>	< <i>*teutæm</i>		cf. Gaul. <i>-an</i> , <i>-im</i> , Lep. <i>-a(:)m</i> , Clb. <i>-a(:)m</i>
gen.	<i>túaithe</i>	< <i>*teut(i)yās</i>		cf. Gaul. <i>-ās</i> , <i>-iās</i> , Clb. <i>-āś</i>
dat.	<i>túait<sup>L</sup></i>	< <i>*teutī</i>		cf. Gaul. <i>-ai</i> , <i>-ī</i> , Lep. <i>-ai</i> , Clb. <i>-ai</i>
<i>Plural</i>				
nom./voc.	<i>túatha</i>	< <i>*teutās</i>		cf. Gaul. <i>-ās</i> (?), Clb. <i>-āś</i>
acc.	<i>túatha</i>	< <i>*teutā(n)s</i>		cf. Gaul. <i>-ās</i> , Clb. <i>-āś</i>
gen.	<i>tuath<sup>N</sup></i>	< <i>*teutōn</i>		cf. Gaul. <i>-ānom</i> , Clb. <i>?-āūm</i>
dat.	<i>tuathab</i>	< <i>*teutābis</i>		cf. Gaul. <i>-ābo</i> , <i>-ābi</i>
<hr/>				
nom.	<i>ben<sup>L</sup></i>	< <i>*benā</i>	< <i>*benā</i>	← <i>*g<sup>w</sup>énh<sub>2</sub></i>
acc.	<i>bein<sup>N</sup></i>	< <i>*benæm?</i>	< <i>*benam</i>	< <i>*g<sup>w</sup>énh<sub>2</sub>m</i>
gen.	<i>mná</i>	< <i>*mnās</i>	< <i>*bnās</i>	< <i>*g<sup>w</sup>néh<sub>2</sub>s</i>
dat.	<i>mnat<sup>L</sup></i>	< <i>*mnāi</i>	< <i>*bnāi</i>	< <i>*g<sup>w</sup>néh<sub>2</sub>(e)i</i>

In Old Irish there is also *bé<sup>N</sup>* ‘woman’, and it has been suggested that this derives < *\*ben* < *\*g<sup>w</sup>ēn* < *\*g<sup>w</sup>ēn* < *\*g<sup>w</sup>énh<sub>2</sub>*, whereas *ben* < *\*benā* (or *\*benā*) is analogical (Jasanoff 1989). This depends on the doctrines that IE */-VRH/* > IE */-V:R/*, and that CC */V:/* > CC */V/* before */-m/* and */-n/* (p. 368).

The only other Old Irish paradigm which may preserve ablaut in the oblique cases is the *n*-stem *cú* (= Brythonic *ki* < *\*kū*), in which oblique *con-* may partly continue *\*kwon-* and may partly represent the weak grade *\*kun-* with regular lowering of */u/* before following */o/* (see Joseph 1990, Stüber 1998: 85–89, de Bernardo Stempel 1999: 27–28, and Table 7.9).

**TABLE 7.9 NOMINAL STEM IN *-N-***

<i>Singular</i>					
nom.	<i>cú<sup>L</sup></i>	< <i>*kū</i>	< <i>*k(u)wō</i>	cf. Skr.	<i>śvā</i>
acc.	<i>coin<sup>N</sup></i>	< <i>*konœn</i> ?	< <i>*kwonm</i>	cf. Skr.	<i>śvānam</i>
gen.	<i>con</i>	< <i>*kunos</i>	< <i>*kunos</i>	cf. Skr.	<i>śúnas</i>
dat.	<i>coin<sup>L</sup></i>	(from acc.)	( <i>*kunej</i> , <i>-i</i> )	cf. Skr.	<i>śúnā</i> (instr.)
<i>Plural</i>					
nom.	<i>coin</i>	< <i>*kones</i>	< <i>*kwones</i>	cf. Skr.	<i>śvānas</i>
acc.	<i>cona</i>	< <i>*kunās</i>	< <i>*kunms</i>	cf. Skr.	<i>śúnas</i>
gen.	<i>con<sup>N</sup></i>	< <i>*kunōn</i>	< <i>*kunōm</i>	cf. Skr.	<i>śúnām</i>
dat.	<i>conaib</i>	< <i>*kunobis</i>	< <i>*kwñb<sup>h</sup>is</i>	cf. Skr.	<i>śvábhiṣ</i> (instr.)

Most other IE declensional classes, for example, *i*-stems, *u*-stems, various consonantal stems and even heteroclitic *r*-/*n*- stems (Stüber 1998: 84, de Bernardo Stempel 1999: 130–139), are represented in Insular Celtic, but much less completely on the Continent; for reasons of space they are omitted here. Adjectives belong to a more restricted range of declensions, mostly vocalic stems. Celtic retained the IE degrees of comparison – the comparative, mostly in *-(i)u* in Old Irish < \**-yōs*, e.g. *siniu* ‘older’ (: Lat. *senior*), also residually in Brythonic, e.g. *W hŷn* ‘older’ < \**senyōs* (de Bernardo Stempel 1989, Schrijver 2007), and the superlative, in *-em*, *-am* in Old Irish, *-ham* in Old Welsh, < \**-isamo/ā* (Gaul. *Marti Rigisamo*, cf. **Position within IE**, p. 357). Celtic also added the equative, in *-ithir*, *-idir* in Old Irish, but *-(h)et* in Brythonic; its etymology is uncertain (cf. Watkins 1966: 37, McCone 1994: 125 = 2005: 141–142).

### Pronouns

Pronouns are not well attested in Continental Celtic, and in Insular Celtic they have evolved through analogical levelling, through interaction with verbal endings and through phonetic attrition, especially when unstressed (see Schrijver 1997, Katz 1998).

The demonstratives mostly derive from IE \**so* + clitic pronouns, e.g. Gaul. *sosin celicnon* (acc.), *sosio* < ?\**sosyod* ‘this’ (p. 361; **Syntax**, p. 377), OIr. *suide* < \**so-de-so(s)* according to Schrijver (1997: 33), not < \**sodyo-*. Demonstratives recalling Lat. *iste* occur in Lep. *isos* (p. 359) and in Clb. *išTe*, *šTena* (n. pl.) (Eska 1989: 165, Schrijver 1997: 63). In Insular Celtic the definite article comes from \**sindos*, \**sindā*, \**sen*; cf. Gaul. *indas mnas* = OIr. *inna mná* ‘the women’ (acc. pl.). Demonstrative \**so* or \**son* may mark the relative in the Old Irish third-person singular relative *beires* ‘who carries, which he carries’ < ?\**beret-so-*, while *(s)a<sup>N</sup>* < \**sen* appears as antecedent in, for example, OIr. *for(s)a<sup>N</sup>* ‘on which’; but the usual relative marker was uninflected \**yo* (: Hitt. *ya* ‘and’): for example, \**esti-yo* ‘who is’ > OIr. *as(as)*, MW *yssyd*; \**welesi-yo* ‘\*whom thou seest’ (Sims-Williams 1984: 153–154) > OIr. *file* ‘who is’; \**beronti-yo* ‘who carry, which they carry’ > OIr. *berte*; cf. Gaul. *dugiiontiio* (p. 361). In compound verbs \**yo* was infix, causing lenition, e.g. *do-ceil* ‘hides’, relative *do-cheil* < \**di-yo-kelet(i)* (McCone 1980, Ó hUiginn 1986). The interrogative stem \**k<sup>w</sup>ey-* (OIr. *cia*, OW *pui* ‘who?’) is rarely used to express the relative. The connective particles \**k<sup>w</sup>e* (: Lat. *-que*, Clb. *-Cue*) > OIr. *-ch* (‘and’) and \**de* (: Gr. *δέ*) > OIr. *-d-* may serve as relative markers in Old Irish, by a secondary development (Vendryes 1911, Watkins 1963).

In the personal pronouns the distinction between nominative and accusative in the first and second person seems to have been lost in Celtic, and in Insular Celtic “dative” pronouns occur, in greatly reduced form, only in combination with other words, for example, \**tī* < \**toy* in OIr. *duit* ‘to you’. First-person singular OIr. *mé* and OBret. *me* suggest \**mē*, whereas OW *mī*, if not influenced by second-person singular *tī*, may derive from \**mī* < \**mē* (cf. *mī* in Gaulish, below). OIr. gen. *mo<sup>L</sup>* ‘my’, stressed *muí* ‘mine’, imply \**mowe* < \**mewe*, whereas MW *vy<sup>N</sup>* ‘my’ implies \**men* < \**mene* (cf. Av. *mana*, Slav. *mene*), so the Old Irish forms (and MW stressed *meu* ‘mine’) may be analogical to the second-person singular \**tewe*. The Old Irish second-person singular *tú* and OW *tī* together imply CC \**tū*. Its genitive forms, OIr. *do<sup>L</sup>*, stressed *tuí*, *taí*, and MW *dy<sup>L</sup>*, stressed *teu*, imply Celt. \**tewe* (: Skr. *táva*). The first-person plural forms, OIr. *sní*, MW *nī*, Gaul. *sni*, suggest \**snī(s)* < \*(s)ne(:)(s), and the second-person plural forms, OIr. *si*, MW *chwi*, suggest \**swī(s)* < \*(s)w(e:)(s); the \**s-* may be due to the first-person plural verbal ending: \**-mos nī(s)* > \**-mos snī(s)* (cf. Katz 1998). The Old Irish genitive forms *nathar*, *nár* ‘ours’ (unstressed *ar<sup>N</sup>*, MW *an*), *sethar*, *sár* ‘yours’ (unstressed *far<sup>N</sup>*), can be compared with Lat. *noster*, *vester*, Goth. *unsara-*, *izwara-* etc. (Schrijver 1995: 454).

The third-person pronouns are problematic. The Old Irish subject pronouns, singular *é* (m.), *sí* (f.) (W *hi*), *ed* (n.), plural *é* (W *wy*), may derive from \*es, \*sī, \*edā, \*ey. W *ef* ‘he’ may come from accusative \*emem, while an unreduplicated accusative \*em lies behind the masculine infixed object pronouns OIr. -a<sup>N</sup>-, Bret. -en-. Old Irish accusative feminine -s<sup>N</sup>-, -e<sup>N</sup>-, may imply \*(s)iyam or \*sām, and the neuter accusative singular -a<sup>L</sup>- implies \*e < \*ed. (In Gaulish *id* may occur, but OIr. *beirthe* ‘carries it’ implies \*bereti-e(d), not \*\*bereti-id.) The accusative plural -s- (also -s<sup>N</sup>-) comes from \*sūs < \*sō(n)s (cf. Gaul. *sos* ‘them’). Most of the above forms cannot be traced *directly* back to IE. By contrast, the unstressed genitive forms OIr. a<sup>L</sup> (m. n.) (MW *y<sup>L</sup>*), a (f.) (MW *y<sup>S</sup>*), a<sup>N</sup> (pl.) (W *eu h-*) can be derived regularly from IE \*esyō, \*esyās, \*eysōm; the /y/ survives as /ð/ in the stem of the Middle Welsh stressed forms *eidaw* (m.), *eidí* (f.).

Subject pronouns seem to occur after verbs in Gaulish (*uediū mi Consonantism*, p. 363, unless -*mi* is an added athematic ending as in Skr. *bhārāmi*), and they perhaps lie behind OIr. 1 sg. -*mm* in, for example, *benaimm* ‘I strike’ < ?\*binam-me. Some Old Irish “emphasizing pronouns” are personal pronouns, for example, *laimir-sni* ‘we dare’ and also (but see Griffith 2010) *ní-bir-siu* ‘you do not carry’ < \*nīs-beres-tū (cf. MW *kereist* ‘you loved’ with -*t* < \*tī); most of them, however, are demonstratives in origin, for example, *beirid-som* ‘he carries’ (: Goth. *sama* ‘the same’, Gr. ὁμός), earlier -*sa* < ?\*se. Old Irish object pronouns are suffixed to simple verbs, for example, *beirthe* ‘carries it’ < \*bereti-e (above), but infixed within compound verbs and after particles, for example, *da-chèil* ‘hides it’ < \*di-e-kelet(i), *ra-ngáid* ‘has beseeched him’ < \*pro-em-g<sup>wh</sup>ōd<sup>b</sup>e. They are often combined with the particle \*de, for example, *fórdom-chàin* ‘he teaches me’ < \*wer-de-me-kanet(i). The Old Irish stress (here indicated by `) regularly follows infixes, and the latter cause sandhi mutations. Deuterotonic compound verbs without visible infixes or sandhi, for example, *do-cèil* ‘hides’, have been attributed either to a meaningless sandhi-inhibiting particle, for example, \*di-(e)s-kelet(i) (Cowgill 1985; for other suggested particles such as \*eti see Karl & Stifter 2007: 301–402), or to analogical creation by “infix deletion”, for example, \*di-e-kelet(i) > \*d(i)-e-xèle(p) → \*di-kèle(p) > *do-cèil* (McCone 1985 – the -*c-* is not lenited because it arose after the period of lenition, p. 367, *pace* Russell 1995: 54). The sentence-initial position of such deuterotonic verbs in Old Irish (**Syntax**, p. 377) implies the former presence of infixes, standing in second position according to Wackernagel’s Law (Vendryes 1911; Chapter 2, p. 377). The deleted particles may have been mainly proleptic/redundant neuter pronouns (e<sup>L</sup>, d(e)-e<sup>L</sup>), whose deletion would have avoided confusion with relative verbs of the type *do-chèil* (p. 370) (Sims-Williams 1984; for other theories see Karl & Stifter 2007: 301–402). For *prototonic* verbs (e.g. *dichil* < \*di-kelet(i)) see **Syntax**, p. 377; such forms occur when a proclitic particle precedes, with or without an infix, e.g. *ní-dichil* ‘does not hide’, *ním-dichil* ‘does not hide me’.

### The Celtic verbal system

This simplified semantic distinctions carried by inflections in Indo-European. For example, the IE aorist and perfect merged in a single “preterite” tense, and the subjunctive and optative moods merged as a single “subjunctive” mood; instead, aspectual differences were expressed syntactically by the use of preverbs and particles, for example, OIr. *ro*, MW *ry* < \*pro (Schmidt 1990). The inflectional system survives most fully in Old Irish. In Old Irish, verbs express the active *voice* with either active *inflection* or deponent *inflection*, and for the passive/impersonal *voice* a passive *inflection* obtained which was similar to but distinct from the deponent inflection; e.g. *suidigidir* ‘places’ differs from *suidigthir* ‘is placed’ only in lacking syncope. According to one theory,



both resulted from a late split in the IE mediopassive: syncope patterns would vary according to the syllable-count of the base, and this variation could have been exploited in order to differentiate deponent and passive (McCone 1986: 240; 2006: 137–138; see differently Jasanoff 1997, Griffith 2009). Already in Gaulish, deponent verbs (i.e. verbs with active meaning and mediopassive inflection) are apparent, e.g. *marcosior* ‘I shall (or may I) ride’ (Lejeune 1985a: 138, Lambert 1994: 125, Delamarre 2003: 217–218). For a detailed lexicon of Celtic verbs see Schumacher 2004, supplemented for Middle Welsh by Rodway 2013.

### *Personal endings*

The personal endings of the Celtic verb derive mainly from:

- (a) the primary endings of the IE present/aorist system (\*-ō/\*-mi, \*-si, \*-ti, etc.), which probably fell together with the secondary endings (\*-m, \*-s, \*-t, etc., Chapter 4, p. xx), partly through loss of \*-i as in Italic (Cowgill 1975, 1985, Lambert 1994: 63, Villar 1995, McCone 2006: 106, Sims-Williams 2007b: 29 and 34) and possibly through a still earlier expansion of the domain of primary endings;
- (b) IE imperative endings;
- (c) IE mediopassive endings in *-r* (as in Hittite, Italic, Tocharian: cf. Chapter 2, p. 99);
- (d) IE perfect endings (see Chapter 2, p. 98–99).

In Insular Celtic there are also obscure “imperfect” endings in the imperfect indicative, conditional and past subjunctive. They are identical in both active and deponent verbs and are apparently of mediopassive origin (cf. Schrijver 1992, Ahlqvist 1993).

Like other IE languages, Celtic came to prefer “thematic” inflection, with a thematic vowel alternating between *e* and *o* before the personal endings, to “athematic” inflection, without *e/o* but often with ablaut variation in the root in the present (full grade in the singular, weak grade in the plural). A survivor of the athematic type is the “copula” (the form of the verb ‘to be’ expressing equivalence rather than existence): third-person singular OIr. *is*, MW *ys* < \*ēs-ti, third-person plural OIr. *it*, MW *ynt* < \*s-énti. Celtic tended to thematize athematic verbs and to generalize a single grade of ablaut, usually the zero grade of the plural: thus IE \*mélǵ-ti, \*mǵ-énti ‘milk(s)’ → CC \*mlig-e-ti, \*mlig-o-nti > OIr. *mligid*, *mlegait* (Watkins 1962: 141–142, McCone 1986: 228, 1991: 29). The \*e/o spread wherever it helped to avoid awkward consonant clusters; hence, it was not inserted after roots with a final laryngeal, which gave Celtic /a/ and remained athematic, e.g. \*skérH-ti, \*skṛH-énti > *skarati*, \*skaranti > OIr. *scaraid*, *scarait* ‘separate(s)’ (cf. Watkins 1962: 189).

- (a) *Primary endings* occur in Old Irish active verbs in the present indicative, present subjunctive and future indicative, and in such preterite indicatives as derive from the IE aorist (i.e. not those which derive from the IE perfect). In Insular Celtic this type of ending has two forms: (a) “absolute”, in simple verbs in absolute initial position without a preceding particle; and (b) “conjunct”, in all compound verbs and in simple verbs in non-initial position: e.g. (a) OIr. *beirid* ‘carries’, MW *trenghit* ‘dies’; (b) *ní-beir* ‘does not carry’, *ny threingk* ‘does not die’. It is now agreed that “absolute” and “conjunct” endings have a single origin, and do not derive from the

IE primary and secondary endings respectively (see Sims-Williams 1984, Cowgill 1985, McCone 1985, 2006; cf. Watkins 1963). Accepting an early apocope of *\*-ī* (p. 372, cf. *sistat* p. 360), the *conjunct* forms of *beirid* ‘carries’ can be derived from primary forms as follows:

**TABLE 7.10 A POSSIBLE ORIGIN OF OLD IRISH CONJUNCT ENDINGS**

1 sg.	<i>-biur</i> [bʲiur]	< *birū	< *bʰerō
2 sg.	<i>-bir</i> [bʲirʲ]	< *biris	< *bʰeresi
3 sg.	<i>-beir</i> [bʲerʲ]	< *beret	< *bʰereti
1 pl.	<i>-beram</i> [bʲerəm]	< *beromos	< *bʰeromosi
2 pl.	<i>-beirid</i> [bʲerʲəðʲ]	< *berete	< *bʰerete
3 pl.	<i>-berat</i> [bʲerət]	< *beront	< *bʰeronti

The Old Irish *absolute* forms are longer than the *conjunct* forms, and it appears that they originally included some additional element which protected final *\*-i* from apocope. Since absolute verbs always occur at the head of clauses, the position of the additional element must be due to Wackernagel’s Law (cf. p. 371). Various particles have been proposed (cf. p. 371), but one candidate (Sims-Williams 1984) is a redundant or proleptic neuter object pronoun *\*e(d)*:

**TABLE 7.11 A POSSIBLE ORIGIN OF OLD IRISH ABSOLUTE ENDINGS**

1 sg.	<i>biru</i> [bʲiru]	< *birū-e
2 sg.	<i>biri</i> [bʲiri]	< *birisi-e
3 sg.	<i>beirid</i> [bʲerʲəðʲ]	< *bereti-
1 pl.	<i>bermai</i> [bʲermi]	< *beromosi-e
2 pl.	<i>beirthe</i> [bʲerʲʰe]	< *berete-e
3 pl.	<i>berait</i> [bʲerədʲ]	< *beronti-

In the third persons – or all persons, according to McCone (1994: 141, 2006: 118, 120, 124), who posits 2 pl. *\*-tesi* and analogical developments in the first-person singular – the additional element was dropped before the general apocope by a process of “suffix deletion” presumably contemporary with “infix deletion” in compound verbs (p. 371).

- (b) The *imperative endings* were mostly similar to the *conjunct* of the present indicative, except the second-person singular, e.g. OIr. *gaib* ‘get!’, Gaul. *gabi*, and the third-person singular, e.g. OIr. *gaibed* < *\*gʰabʰyetow?* (cf. Goth. *-dau*); the need to differentiate the indicative from the imperative in absolute initial position may have encouraged the generalization in that position of the absolute (rather than *conjunct*) indicative endings of simple verbs (Sims-Williams 1984: 171, Eska 1991). Similarly, the imperative of compound verbs was distinguished from the (deuterotonic) indicative by its protonic stress (cf. p. 371).
- (c) The *passive/impersonal* inflection has third-person singular and third-person plural endings only. (The absolute forms, with palatalized final consonant, are arguably based on the analogy of active absolute forms like third-person plural *berait*.) In the present tense there are two inflections in the singular, with or without a dental

consonant (cf. Umbr. pres. subj. *ferar* vs Lat. *feratur?*), as seen in the following conjunct forms:

TABLE 7.12 OLD IRISH PASSIVE FORMS (PRESENT TENSE)

3 sg.	<i>-berar</i>	< *beror	(absolute <i>berair</i> ) ‘is carried’
	<i>-marbthar</i>	< *marwa(:)tor	(absolute <i>marbthair</i> ) ‘is killed’
3 pl.	<i>-bertar</i>	< *berontor	(absolute <i>bertair</i> )
	<i>-marbtar</i>	< *marwa(:)ntor	(absolute <i>marbtair</i> )

Infixed pronouns indicate the first and second persons, e.g. *nom-berar* ‘I am carried’.

The passive preterite paradigm, however, was built up on the basis of the IE verbal adjective (Sims-Williams 1984: 183), e.g. MW *llas* ‘he was slain’ < \*slad-tos, OIr. *nom-breth* ‘I was carried’ < \*nu-me-britos (< \*b<sup>h</sup> ʔtos).

The *deponent* conjunct endings may be illustrated with OIr. *-fograigedar* ‘sounds’. The irregular syncope outside the third-person singular and third-person plural of this five-syllable base arguably follows the pattern of four-syllable bases (p. 372). (Jasanoff (1997), however, hypothesizes 3 sg. \*-tro, 3 pl. \*-ntro.)

TABLE 7.13 OLD IRISH DEPONENT FORMS (PRESENT TENSE)

1 sg.	<i>-fograigir</i>	< *wogaro-sagī-ōr	(absolute <i>fograigim</i> )
2 sg.	<i>-fograigther</i>	< *wogaro-sagī-ter	(absolute <i>fograigther</i> )
3 sg.	<i>-fograigether</i>	< *wogaro-sagī-tor	(absolute <i>fograigidir</i> )
1 pl.	<i>-fograigmer</i>	< *wogaro-sagī-mor	(absolute <i>fograigmir</i> )
2 pl.	<i>-fograigid</i>	< *wogaro-sagī-dwe (?)	(absolute <i>fograigthe</i> )
3 pl.	<i>-fograigetar</i>	< *wogaro-sagī-ntor	(absolute <i>fograigitir</i> )

The first-person singular and second-person plural absolute endings are borrowed from the athematic active, and the palatalized consonants in the absolute third-person singular, first-person plural and third-person plural are by analogy with passive absolute forms like *gaibthir*, *gaibtir*.

- (d) Some IE *perfect endings* survived in the OIr. ‘suffixless preterite’, e.g. 1 sg. *-gád* ‘I prayed’ < \*g<sup>w</sup>āda, 3 sg. *-gáid* < \*g<sup>w</sup>āde, 3 pl. *-gádatar* < \*g<sup>w</sup>ādontVr – the last a blend of the IE *r*-ending (cf. Skr. *vid-úr*) and the *nt*-ending of the present/aorist (cf. Lat. *uidē-r-unt*).

### Present stem

The present stem (used to form the present and imperfect indicative and the imperative) was formed with various suffixes, which merged to give the following Old Irish conjugations according to the numeration of Thurneysen 1946:

- AI: \**-ā-* (cf. Lat. *-are*) and \**-ǎ-* < root-final \*-H, e.g. *scaraid* (**Personal Endings**, p. 372);  
 AII: \**-eye/eyo-* (denominative and causative), also \**-ī-* (partly < stative \**-ē-*, e.g. *-ruidi* ‘blushes’ cf. Lat. *rubēre* ‘to be red, blush’);

- AIII: miscellaneous verbs with hiatus, e.g. *baid* ‘dies’ < \*ba-eti (McCone 1986: 228);  
 BI: \*-e/o- (e.g. *beirid*, **Personal Endings**, p. 372);  
 BII: \*-ye/yo- (3 sg. *-gaib*, not \*\**-gaibi* < \*g<sup>h</sup>ab<sup>h</sup>yet(i), is arguably on the analogy of BI, although a \*-ī/yo- suffix or \*ye > \*i have been suggested, cf. Sims-Williams 1981: 211–216, Weiss 2012);  
 BIII: \*-e/o- with nasal infix, e.g. *bongid* ‘reaps’, passive pret. *-bocht*, cf. Lat. *ta-n-go, tac-tus* (cf. Joseph 1990, McCone 1991: 41–47);  
 BIV: \*-nā-, e.g. *crenaid*, pl. *crenait* ‘buys’ < \*k<sup>w</sup>ri-nā-ti, \*k<sup>w</sup>ri-nā-nti (? ultimately from \*k<sup>w</sup>ri-né-h<sub>2</sub>-ti, \*k<sup>w</sup>ri-n-h<sub>2</sub>-énti, with nasal infix in the root \*k<sup>w</sup>r(e)ih<sub>2</sub>-, seen without infix in the Old Irish subjunctive stem *cria-*) (McCone 1986: 225; cf. 1991: 11–54);  
 BV: supposedly \*-n(e)u- (but see McCone 1986: 225–227, Campanile 1990, Hamp 1991, McCone 1991: 13–15, 22–23).

AI and AII are the only productive formations in Old Irish, and broadly resemble the Brythonic regular verbs, e.g. MW *caraf* ‘I love’ < \*karāmi, *kenif* ‘I sing’ < \*kanīmi (unless the latter is < \*kanū-mi).

### Subjunctive stem

The subjunctive stem (used for the present and past subjunctive) is divided into two classes: (a) the unproductive *s*-subjunctive, e.g. OIr. *geiss*, *-gé* ‘may pray’ < \*g<sup>w</sup>ed-s-t(i), MW *gwnech* ‘may do’ < \*(g)wrex < \*wrek-s-et(i); and (b) the productive so-called *ā*-subjunctive, e.g. *beraid*, *-bera* ‘may carry’, traditionally derived from \*ber-ā-t(i). It is semantically difficult to derive (a) from the IE *s*-aorist indicative (with Watkins 1962), and derivation from the *s*-aorist subjunctive is difficult unless \*g<sup>w</sup>ed-s-t(i) replaced the expected \*g<sup>w</sup>ed-s-et(i) under the influence of the *s*-preterite inflection (McCone 1986: 245–246), which is against the normal Celtic athematic → thematic tendency (cf. Hamp 1987, McCone 1991: 57, 73, 79–80, Schumacher 2004: 53). Subjunctives in *-se-Ti* may occur in Celtiberian (Eska 1989: 170). (b) The *ā*-subjunctive, traditionally derived, with that of Latin, from an innovative Italo-Celtic optative suffix \*-ā- (cf. OLat. *aduenat*, subj. of *aduenio*), is now often analysed as \*-ā-se/o- or \*-ā-se/o-, with \*ā, \*ā originally from roots in \*-H, \*-R<sub>H</sub> (McCone 1986: 260, 1991: 85–113, Schumacher 2004: 50); other linguists retain the Italo-Celtic explanation (cf. Schmidt 1991: 17–19; Jasanoff 1994).

### Future stem

There are three main types of Old Irish future stem (used for the indicative future and conditional): (a) the *f*-future, e.g. *rannfa*, *-rannub* ‘I shall divide’; (b) the reduplicated *s*-future, e.g. *gigis*, *-gig* ‘will pray’ < \*g<sup>w</sup>i-g<sup>w</sup>ed-s-t(i); and (c) the reduplicated so-called *ā*-future, e.g. *cechnaid*, *-cechna* ‘will sing’ < \*ki-kana(:)-se-t(i), and its subtype, the *ē*-future, e.g. *célaid*, *-céla* ‘will hide’ < ?\*ki-kla(:)-se-t(i) (with \*kikl- > \*kexl- > *cēl-*). (a), the *f*-future, is traditionally compared with the Latin future in *-bo*. This is phonetically controversial – normally /f/ is < \*sw – but because the *f*-future occurs only in Old Irish, where it is very productive, it is tempting to see it as a late, parallel innovation, based on a periphrasis involving the root \*b<sup>h</sup>(e)w- ‘to be’ (but on /bw/ cf. McManus 1991: 122). However, McCone (1991: 176–182) suggested that it spread from a verb in which reduplicated \*si-sw- > \*si-f-; Jasanoff (1994: 216) and Russell (1995: 20 and 49) rejected this. (b–c), the two reduplicated futures, run parallel to the corresponding subjunctives, and

their suffixes must be explained similarly. They are originally desideratives (semantically cf. English ‘he *will* pray’) and are comparable with Sanskrit desideratives, e.g. *tīṭṭṣati* < \*tī-ṭṭp-se-ti (root \*terp- ‘enjoy’) (Thurneysen 1946: 414–415; McCone 1986: 248–255). Unreduplicated futures in \*-sye/o- have been identified in Continental Celtic, Indo-Iranian and elsewhere (e.g. Schmidt 1988: 241, Lambert 1994: 63, Schumacher 2004: 58).

### *Active preterite stem*

The Old Irish active preterite stem (to which “deponent” as well as “active” inflections could be added) derived, for most verbs, from the IE aorist, but for others from the IE perfect.

With few exceptions the aorists were originally athematic and sigmatic. According to Watkins (1962) third-person singular forms, e.g. \*skerH-s-t, \*ber-s-t developed regularly via \*skarass(i), \*bert(i) to *-scar* ‘separated’, *-bert* ‘carried’ (cf. OIr. *tart* < \*tṛstu- ‘thirst’), and these third-person singular forms formed the basis for the whole paradigms, e.g. first-person singular \*skarass-ū > *-scarus* (the *s*-pret.), \*bert-ū > *-biurt* (the *t*-pret.). Cf. MW 1 sg. *kereis* ‘loved’ < \*karassū, 3 sg. (\*)*caras* < \*karassit(i); other stem vowels before \*ss were partly influenced by present stems (McCone 1986: 232, Rodway 2013: 69–80, 128–165). A variant on Watkins’ *s*-aorist theory derives *-bert* < \*birt < \*bīr(s)t(i) < \*bīrst(i), with lengthened grade; this vocalism is supported by the \*byrth underlying the MW compound *kymyrth* ‘took’ (McCone 1991: 67–69; Schumacher 2004: 61–64). It is phonologically more difficult to apply Watkins’ *s*-aorist theory to a Middle Welsh *t*-preterite such as MW 1 sg. *keint* ‘sang’ < \*kantū, 3 sg. *cant* < \*kant(i), and similarly 3 sg. *gwant* ‘pierced’. Isaac (2001) explained the latter as an IE athematic imperfect, and Jasanoff (2012), accepting this, has explained MW *kymyrth*, OIr. *-bert* and other *t*-preterites as IE imperfects with lengthened grade \*ē.

The third-person singular Brythonic termination of Middle Welsh absolute *keressyt* < \*karass-iti, conjunct (\*)*caras* < \*karass-it, is an innovation paralleled in Gaul. *legasit* < ?\*legast + \*-iti versus more archaic *prinas* < \*k<sup>w</sup>rinast (Sims-Williams 1984: 188; cf. Lambert 1994: 64, 68, Eska & Evans 2010: 41). For other Continental Celtic preterites see p. 359.

Some of the Old Irish “suffixless preterites” employ reduplication and derive from old perfects (cf. Lat. *tango*, *te-tigi*), e.g. OIr. *cechain* ‘he sang’ < \*ke-kan-e; a rare MW example is *kigleu* ‘he heard’. Others replaced \*ē in the root with Celtic \*ā (probably < \*ō), e.g. OIr. *do-feid* < \*-wedet ‘leads’, *do-fāid* < \*-wāde ‘led’, MW *godiwawd* < \*-wāde ‘overtook’. It is unclear whether this *ā*-preterite is attested in Continental Celtic: a possible precursor, perhaps with /ō/ as in the IE perfect, has been seen in Gaul. *AVVOT(E)* ‘*FECIT*’ = ? /aw-wo(:)d(e)/ < \*a(p)jo-wōd<sup>(h)</sup>-e ‘led away, carried out, produced’ (Lambert 1987, 1994: 122; cf. Hamp 1973). In general, the Celtic *ā*-preterite resembles long-vowel preterites in other languages (cf. Germanic comparison in McCone 1986: 235–238, 1994: 168 = 2005: 224) and like them is of obscure origin. Jasanoff (2012) derives them from IE imperfects.

## WORD FORMATION

Despite much work on Celtic name formation (e.g. Uhlich 1993, García Alonso 2013, Falileyev 2014), the vast subject of Celtic word formation has only recently been studied in detail, for example, Joseph 1987 on denominative verbs in \*-sag- ‘seek’ (e.g. *fograige-dar*: *fogur* ‘sound’, **Personal Endings**, p. 374), Russell 1990 on velar suffixes (such as the productive \*-āko-), Irslinger 2002 and Repanšek 2013 on dental suffixes, de Bernardo

Stempel 1999 and Zimmer 2000 on Old Irish and Welsh noun formation, Stüber 1998 on *n*-stems, and Schumacher 2000 on Welsh verbal nouns. Most IE types of composition survive in Celtic, at least residually (cf. McCone 1994: 126–132 = 2005: 146–155), including *dvandva*-compounds such as OIr. *gaisced* ‘weapons’ < \**gaiso-skeitom* ‘spear-and-shield’ or *TEVO-XTONION* (gen. pl.) = *DEIS ET HOMINIBUS* in a bilingual Gaulish inscription at Vercelli (Delamarre 2003: 141–142, 176, cf. p. 368). Composition vowels are discussed by Sims-Williams (2013). Like other IE languages, Celtic developed the use of preverbs to modify verbal bases aspectually (**The Celtic Verbal System**, p. 371) or semantically, e.g. OIr. *fo-reith* ‘helps’ < \*‘runs under’ < *reithid* ‘runs’ (: MW *gwaret* ‘helps’; Lat. *subcurro, succurro* < *curro*). Large numbers may be strung together (e.g. OIr. *intururas* ‘incurSION’ < \**ind-to-are-uks-ret-*), and they tend to appear in a particular hierarchy (McCone 1987: 94) which has parallels elsewhere, e.g. *ro* < \**pro* tends to occur close to the root, as did Ved. *prá*, Hom. *pró* (Sims-Williams 1984: 190). Despite similarities, the system of preverbs and prepositions in Celtic and Italic cannot be reduced to a unity (Watkins 1966: 36). As in other IE languages, e.g. OLat. *ob uos sacro* (→ *obsecro uos*), preverbs may be divided by tmesis (see **Syntax** below; McCone 1985: 267).

## SYNTAX

The most important development in Celtic is in the position of the verb. Whereas the normal order is SOV in Celtiberian (e.g. p. 360) and in some Gaulish (e.g. *Buscilla sosio legasit in Alixie Magalu* ‘B. chose this in Alisia for Magalos’), Insular Celtic favours VSO (p. 368), with most apparent exceptions, such as SVO (Lewis 1989), being explainable either as *nominativus pendens* or as cleft sentences with [copula] + S + relative verb (Russell 1995: 292–300, Willis 1998: 97–101). The most important exception is the arguably archaic Old Irish construction known as Bergin’s Law (Binchy 1979), by which the verb, instead of appearing initially in *absolute* (**Personal Endings**, p. 372–374) or *deuterotonic* (p. 371) form (e.g. \**Loiscis Lugaid trebthu* ‘Lugaid burnt dwellings’, \**Ad-rimi maicni nAilb* ‘You reckon the sons of Alb’), appears finally/medially in *conjunct* or *prototonic* form, e.g. *Lugaid loisc trebthu, Maicni nAilb áirmi*. This construction is attributed by some to preservation of the SOV/SVO syntax seen in Continental Celtic but by others to mere poetic licence (McCone 2006: 61–65, Eska 2007). The key to the development of VSO was identified by Vendryes (1911; cf. Watkins 1963, Eska 1994, Russell 1995: 13, 303, McCone 2006) in the phenomenon that certain clitics were closely tied to verbs in Celtic and therefore drew either the verb or its first preverb to the head of the clause by Wackernagel’s Law (p. 371; **Personal Endings**, p. 372–373), e.g. verb + relative \**yo* in Gaul. *dugiiontiio Ucuetin* ‘who honour Ucuetis’ (p. 361), preverb *imm* + *-a* < \**yo* in archaic/poetic OIr. *imma- lanna -lig* ‘which lies about lands’ (with tmesis (**Word Formation**, p. 377), → later/prosaic \**imma-lig lanna*). This phenomenon (labelled “Vendryes’ Restriction” in Sims-Williams 1984: 167) must surely be linked with the fact that absolute and deuterotonic verbal forms, which are required in VSO order, seem originally to have included clitic elements subject to Wackernagel’s Law (see **Pronouns**, p. 370–371, **Personal Endings**, p. 372–374; for theories about the identity of these clitics cf. McCone 2006 and Karl & Stifter 2007: 301–402).<sup>1</sup>

## FURTHER READING

More detailed surveys are provided by MacAulay 1992 and Ball and Müller 2010, the former concentrating on the modern languages. The only comparative grammar is still

Lewis and Pedersen 1961. Being a revision of a work published in 1909–13, this does not reflect the vast amount of work on Celtic over the past hundred years. Its old-fashioned IE notation has been followed to some extent in this chapter for the convenience of readers. There is no handbook of Continental Celtic as a whole, but Lambert 1994 is a good introduction to Gaulish while Delamarre 2003 is an etymological dictionary. The classic grammar of Old Irish by Thurneysen (1946) still holds the field, but can be supplemented by McCone 1994 (Italian version 2005). Etymology is not covered in the *Dictionary of the Irish Language* (<http://edil.qub.ac.uk/dictionary/search.php>), and the Old Irish etymological dictionary by Vendryes (1959–) still lacks the letters E–L. The standard grammars for early Welsh, Cornish and Breton by Evans 1964, Lewis 1990 and Hemon 1975 can be supplemented by Ternes 2011. Etymology is not the most up-to-date aspect of *Geiriadur Prifysgol Cymru: A Dictionary of the Welsh Language* (<http://www.welsh-dictionary.ac.uk>), which closely follows Pokorny 1959, but it is a convenient starting point, as are Pokorny 1959 and Matasović 2009 for the Celtic languages as a whole.

### FURTHER READING

In revising this chapter I am grateful to the editor and to J.F. Eska and P. Schrijver for helpful criticism.

### REFERENCES

- Ahlqvist, Anders 1993, “The Old Irish imperfect indicative,” in: Brogyanyi, B. & Lipp, R. (eds.), *Comparative-Historical Linguistics: Indo-European and Finno-Ugric. Papers in Honor of Oswald Szemerényi*, vol. III, Benjamins, Amsterdam: 281–289
- Anderson, Cormac 2014, “Consonant quality in Old Irish revisited,” in: Roma, E. & Stifter, D. (eds.), *Linguistic and Philological Studies in Early Irish*, Lewiston, New York & Lampeter: 1–29
- Ball, Martin J., & Müller, Nicole (eds.) 2010, *The Celtic Languages*, 2nd ed, Routledge, London
- Bammesberger, Alfred (ed.) 1988, *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*, C. Winter, Heidelberg
- Binchy, D. A. 1979, “‘Bergin’s Law’,” *Studia Celtica* 14–15: 34–53
- Bird, Norman 1982, *The Distribution of Indo-European Root Morphemes*, Harrassowitz, Wiesbaden
- Blom, Alderik H. 2009, “*lingua gallica, lingua celtica*: Gaulish, Gallo-Latin, or Gallo-Romance?,” *Keltische Forschungen* 4: 7–54
- Broderick, George 2013, “Indo-European and non-Indo-European aspects to the languages and place-names in Britain and Ireland: an overview,” in: Hambro, Cathinka & Widerøe, Lars Ivar (eds.), *Lochlann: Festskrift til Jan Erik Rekdal*, Hermes Academic Publishing, Oslo: 282–314
- Campanile, Enrico 1970, “Sulle isoglosse lessicali celtogermaniche,” *AION* 9: 13–39
- Campanile, Enrico 1990, “A note on the classification of some Old Irish verbs,” *Celtica* 21: 99–103
- Cowgill, Warren 1975, “The origins of the Insular Celtic conjunct and absolute verbal endings,” in: Rix, H. (ed.), *Flexion und Wortbildung*, Reichert, Wiesbaden: 40–70
- Cowgill, Warren 1980, “The etymology of Irish *guidid* and the outcome of \*g<sup>h</sup> in Celtic,” in: Mayrhofer, M., Peters, M. & Pfeiffer, O. (eds.), *Lautgeschichte und Etymologie*, Reichert, Wiesbaden: 49–78

- Cowgill, Warren 1985, "On the origin of the absolute and conjunct verbal inflexion of Old Irish," in: Schlerath, B. (ed.), *Grammatische Kategorien: Funktion und Geschichte*, Reichert, Wiesbaden: 109–118
- de Bernardo Stempel, Patrizia 1987, *Die Vertretung der indogermanischen liquiden und nasalen Sonanten im Keltischen*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck
- de Bernardo Stempel, Patrizia 1989, "Britannischer Komparativ und Konsonantenverdoppelung," *Indogermanische Forschungen* 94: 207–233
- de Bernardo Stempel, Patrizia 1995, "Gaulish accentuation: results and outlook," in: Eska, Gruffydd & Jacobs 1995: 16–32
- de Bernardo Stempel, Patrizia 1999, *Nominale Wortbildung des älteren Irischen*, Niemeyer, Tübingen
- de Bernardo Stempel, Patrizia 2008, "Cib. to Luguei 'hacia Lugus' frente a Luguei 'para Lugus': sintaxis y divinidades en Peñalba de Villastar," *Emerita* 76: 181–196
- de Bernardo Stempel, Patrizia 2010, "La ley del 1<sup>er</sup> Bronce de Botorrita: uso agropecuario de un encinar sagrado," in: Mozota, Francisco Burillo (ed.), *Ritos y mitos*, Centro de Estudios Celtibéricos de Segeda, Daroca: 123–145
- de Hoz, Javier 2007, "The Mediterranean frontier of the Celts and the advent of Celtic writing," *Cambrian Medieval Celtic Studies* 53/54: 1–22
- de Hoz, Javier 2010–, *Historia lingüística de la península ibérica en la antigüedad*, Consejo Superior de Investigaciones Científicas, Madrid
- Delamarre, Xavier 2003, *Dictionnaire de la langue gauloise*, 2nd ed., Errance, Paris
- de Simone, Carlo 1980, "Gallish \*Nemetijos – etruskisch Nemetie," *Zeitschrift für vergleichende Sprachforschung* 94: 198–202
- Eichner, Heiner 1989, "Damals und heute: Probleme der Erschließung des Altkeltischen zu Zeußens Zeit und in der Gegenwart," in: Forssman, B. (ed.), *Erlanger Gedenkfeier für Johann Kaspar Zeuß*, Univ. Bibl., Erlangen: 9–56
- Elsie, Robert W. 1990, "Proto-Brittonic Celtic and dispersion in the Indo-European lexicon," *Ollodagos* 1: 279–321
- Eska, Joseph F. 1989, *Towards an Interpretation of the Hispano-Celtic Inscription of Botorrita*, Institut für Sprachwissenschaft der Universität, Innsbruck
- Eska, Joseph F. 1990, "Syntactic notes on the great inscription of Peñalba de Villastar," *Bulletin of the Board of Celtic Studies* 37: 104–107
- Eska, Joseph F. 1991, "First person emphatic and imperative in Early Irish," *Bulletin of the Board of Celtic Studies* 38: 87–92
- Eska, Joseph F. 1994, "On the crossroads of phonology and syntax: remarks on the origin of Vendryes' Restriction and related matters," *Studia Celtica* 28: 39–62
- Eska, Joseph F. 1995, "Observations on the thematic genitive singular in Lepontic and Hispano-Celtic," in: Eska, Gruffydd & Jacobs 1995: 33–46
- Eska, Joseph F. 1996, "Resyllabification and epenthesis in Hispano-Celtic," *Journal of Celtic Linguistics* 5: 71–89
- Eska, Joseph F. 2003, "On syntax and semantics in Alise-Sainte-Reine (Côte-d'Or), again," *Celtica* 24: 101–120
- Eska, Joseph F. 2007, "Bergin's Rule: Syntactic diachrony and discourse strategy," *Diachronica* 24: 253–278
- Eska, Joseph F. 2007/2008, "Remarks on the 3. plural preterite in -us in Continental Celtic," *Die Sprache* 47: 108–119
- Eska, Joseph F. 2010a, "The emergence of the Celtic languages," in: Ball & Müller (eds), 2010: 22–27



- Eska, Joseph F. 2010b, "Contact and the Celtic languages," in: Hickey, Raymond (ed.), *The Handbook of Language Contact*, Wiley-Blackwell, Malden, MA: 538–549
- Eska, Joseph F. 2013a, "A salvage grammar of Galatian," *Zeitschrift für celtische Philologie* 60: 51–63
- Eska, Joseph F. 2013b, "In defense of Celtic /φ/," in: Cooper, Adam I., Rau, Jeremy & Weiss, Michael (eds.), *Multi nominis grammaticus: Studies in Classical and Indo-European Linguistics in Honor of Alan J. Nussbaum*, Beach Stave Press, Ann Arbor, NY: 32–43
- Eska, Joseph F., & Evans, D. Ellis 2010, "Continental Celtic," in: Ball & Müller (eds.) 2010: 28–54
- Eska, Joseph F., Gruffydd, R. Geraint & Jacobs, Nicolas (eds.) 1995, *Hispano-Gallobrittonica: Essays in Honour of D. Ellis Evans*, University of Wales Press, Cardiff
- Eska, Joseph F., & Mercado, Angelo O. 2005, "Observations on verbal art in ancient Vergiate," *Historische Sprachforschung/Historical Linguistics* 118: 160–184
- Evans, D. Ellis 1967, *Gaulish Personal Names*, Clarendon Press, Oxford
- Evans, D. Ellis 1979, "The labyrinth of Continental Celtic," *Proceedings of the British Academy* 65: 497–538
- Evans, D. Ellis 1981, "Celts and Germans," *Bulletin of the Board of Celtic Studies* 29: 230–255
- Evans, D. Ellis 1983, "Continental Celtic and linguistic reconstruction," in: Mac Eoin, G. (ed.), *Proceedings of the Sixth International Congress of Celtic Studies*, Dublin Institute for Advanced Studies, Dublin: 19–54
- Evans, D. Ellis 1988, "Celtic origins," in: *MacLennan* 1988: 209–222
- Evans, D. Ellis, Griffith, John G., & Jope, E.M. (eds.) 1986, *Proceedings of the Seventh International Congress of Celtic Studies*, Oxbow Books, Oxford
- Evans, D. Simon 1964, *A Grammar of Middle Welsh*, Dublin Institute for Advanced Studies, Dublin
- Falileyev, Alexander 2009, *Celtic Dacia: Personal Names, Place-Names and Ethnic Names of Celtic Origin in Dacia and Scythia Minor*, 2nd ed, CMCS Publications, Aberystwyth
- Falileyev, Alexander 2013, *The Celtic Balkans*, CMCS Publications, Aberystwyth
- Falileyev, Alexander 2014, *In Search of the Eastern Celts: Studies in Geographical Names, their Distribution and Morphology*, Archaeolingua, Budapest
- Falileyev, Alexander, with Gohil, Ashwin E. & Ward, Naomi 2010, *Dictionary of Continental Celtic Place-Names*, CMCS Publications, Aberystwyth
- Fleuriot, Léon 1978, "Brittonique et gaulois durant les premiers siècles de notre ère," in: *Étrennes de septantaine: travaux de linguistique et de grammaire comparée offerts à M. Lejeune*, Klincksieck, Paris: 75–83
- Fleuriot, Léon 1988, "New documents on ancient Celtic and the relationship between Brittonic and Continental Celtic," in: *MacLennan* 1988: 223–230
- García Alonso, Juan Luis (ed.) 2013, *Continental Celtic Word Formation: The Onomastic Data*, Ediciones Universidad de Salamanca, Salamanca
- Greene, David 1966, "The making of Insular Celtic," in: *Proceedings of the Second International Congress of Celtic Studies*, University of Wales Press, Cardiff: 123–136
- Greene, David 1973, "The growth of palatalization in Irish," *Transactions of the Philological Society*: 127–136
- Griffith, Aaron 2009, "The Old Irish deponent suffixless preterite," *Keltische Forschungen* 4: 169–187

- Griffith, Aaron 2010, "Die Etymologie der 1. und 2. Person Singular der altirischen *notae augentes*," *Keltischen Forschungen* 5: 107–119
- Hamp, Eric P. 1958, "Consonant allophones in Proto-Keltic," *Lochlann* 1: 209–217
- Hamp, Eric P. 1965, "Evidence in Celtic," in Winter, W. (ed.), *Evidence for Laryngeals*, Mouton, The Hague: 224–235
- Hamp, Eric P. 1973, "Some *ā*-preterites," *Celtica* 10: 157–159
- Hamp, Eric P. 1987, "The athematic *s*-subjunctive," *Ériu* 38: 201
- Hamp, Eric P. 1991, "Varia," *Celtica* 22: 33–47
- Hamp, Eric P. 2012, "Gaulish ordinals and their history," *Études celtiques* 38: 131–135
- Harvey, Anthony 1984, "Aspects of lenition and spirantization," *Cambridge Medieval Celtic Studies* 8: 87–100
- Harvey, Anthony 1989, "Some significant points of early Insular Celtic orthography," in: Ó Corráin, D., Breatnach, L. & McCone, K.R. (eds.), *Sages, Saints and Storytellers: Celtic Studies in Honour of James Carney*, An Sagart, Maynooth: 56–66
- Hemon, Roparz 1975, *A Historical Morphology and Syntax of Breton*, Dublin Institute for Advanced Studies, Dublin
- Irslinger, Britta Sofie 2002, *Abstrakta mit Dentalsuffixen im Altirischen*, C. Winter, Heidelberg
- Isaac, G. R. 2001, "An Indo-European athematic imperfect in Welsh? Middle Welsh *gwant*," *Studia Celtica* 35: 354–359
- Jackson, Kenneth Hurlstone 1953, *Language and History in Early Britain*, Edinburgh University Press, Edinburgh
- Jackson, Kenneth Hurlstone 1980, "The Pictish language," in: Wainwright, F.T. (ed.), *The Problem of the Picts*, 2nd ed, Melven Press, Perth: 129–166, 173–176
- Jasanoff, Jay H. 1989, "Old Irish *bé* 'woman'," *Ériu* 40: 135–141
- Jasanoff, Jay H. 1994, "The Brittonic subjunctive and future," in: Rasmussen, J.E. (ed.), *In Honorem Holger Pedersen*, Reichert, Wiesbaden: 199–220
- Jasanoff, Jay H. 1997, "An Italo-Celtic isogloss: The 3 pl. mediopassive in *\*-ntro*," in: Adams, Douglas Q. (ed.), *Festschrift for Eric P. Hamp*, Volume I, Institute for the Study of Man, Washington, DC: 146–161
- Jasanoff, Jay H. 2012, "Long-vowel preterites in Indo-European," in: Melchert, Craig H. (ed.), *The Indo-European Verb*, Reichert Verlag, Wiesbaden: 127–135
- Jaskuła, Krzysztof 2014, "A pragmatic approach to Old Irish consonant qualities," *Journal of Celtic Linguistics* 15: 39–51
- Jordán Cólera, Carlos 2004, *Celtibérico*, Universidad de Zaragoza, Zaragoza
- Joseph, Lionel S. 1982, "The treatment of *\*CR̥H-* and the origin of CaRa- in Celtic," *Ériu* 33: 31–57
- Joseph, Lionel S. 1987, "The origin of the Celtic denominatives in *\*-sag-*," in: Watkins, C. (ed.), *Studies in Memory of Warren Cowgill*, de Gruyter, Berlin: 113–159
- Joseph, Lionel S. 1990, "Old Irish *cú*: A naïve reinterpretation," in: Matonis, A. T. E. & Melia, Daniel F. (eds.), *Celtic Language, Celtic Culture: A Festschrift for Eric P. Hamp*, Ford & Baillie, Van Nuys, CA: 110–130
- Karl, Raimund, & Stifter, David (eds.) 2007, *The Celtic World: Critical Concepts in Historical Studies*, Vol. IV, *Celtic Linguistics*, Routledge, London–New York
- Katz, Joshua T. 1998, "Archaische keltische Personalpronomina aus indogermanischer Sicht," in: Meid, Wolfgang (ed.), *Sprache und Kultur der Indogermanen*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck: 265–291
- Koch, John T. 1987, "Prosody and the Old Celtic verbal complex," *Ériu* 38: 143–176

- Köderitzsch, Rolf 1985, "Die große Felsinschrift von Peñalba de Villastar," in: Ölberg, H.M., Bothien, H. & Schmidt, G. (eds.), *Sprachwissenschaftliche Forschungen: Festschrift für Johann Knobloch*, Institut für Sprachwissenschaft der Universität, Innsbruck: 211–222
- Kuryłowicz, Jerzy 1960, *Esquisses linguistiques*, PAN, Wrocław–Cracow
- Kuryłowicz, Jerzy 1964, *The Inflectional Categories of Indo-European*, C. Winter, Heidelberg
- Lambert, Pierre-Yves 1987, "Notes linguistiques gauloises," in: *Mélanges offerts au Docteur J.-B. Colbert de Beaulieu*, Le Léopard d'Or, Paris: 527–534
- Lambert, Pierre-Yves 1994, *La langue gauloise*, Errance, Paris
- Lambert, Pierre-Yves 2002, *Recueil des inscriptions gauloises, II.2, Textes gallo-latins sur instrumentum*, CNRS, Paris
- Lambert, Pierre-Yves 2010, "La reconstruction du gaulois entre typologie et comparaison," in: Choi-Jonin, Injoo, Duval, Marc & Soutet, Olivier (eds.), *Typologie et comparatisme*, Peeters, Leuven: 265–278
- Lambert, Pierre-Yves, & Pinault, Georges-Jean (eds.) 2007, *Gaulois et celtique continental*, Droz, Geneva
- Lejeune, Michel 1970, "Documents gaulois et para-gaulois de Cisalpine: Lepontica," *Études celtiques* 12: 357–500
- Lejeune, Michel (ed.) 1985a, "Le plomb de Larzac," *Études celtiques* 22: 95–177
- Lejeune, Michel 1985b, "La première déclinaison celtique," *Études celtiques* 22: 88–93
- Lewis, Henry 1989, *Die kymrische Sprache*, Institut für Sprachwissenschaft der Universität, Innsbruck
- Lewis, Henry 1990, *Handbuch des Mittelkornischen* (transl. Stefan Zimmer), Institut für Sprachwissenschaft der Universität, Innsbruck
- Lewis, Henry & Pedersen, Holger 1961, *A Concise Comparative Celtic Grammar*, 2nd edn, Vandenhoeck & Ruprecht, Göttingen
- Lindeman, Fredrik Otto 1988, "Some remarks on 'laryngeals' and Celtic," in: Jazayery, M. A. & Winter, W. (eds.), *Languages and Cultures: Studies in Honor of Edgar C. Polomé*, de Gruyter, Berlin: 397–400
- Lindeman, Fredrik Otto 1989, Review of Bammesberger 1988, in: *Zeitschrift für vergleichende Sprachforschung* 102: 268–297
- Lindeman, Fredrik Otto 1992, "L'accent indo-européen et le verbe celtique," *Études celtiques* 29: 43–49
- MacAulay, Donald (ed.) 1992, *The Celtic Languages*, Cambridge University Press, Cambridge
- MacLennan, Gordon W. (ed.) 1988, *Proceedings of the First North American Congress of Celtic Studies*, Chair of Celtic Studies, Ottawa
- Mallory, J. P. 1989, *In Search of the Indo-Europeans: Language, Archaeology and Myth*, Thames & Hudson, London
- Martinet, André 1952, "Celtic lenition and western Romance consonants," *Language* 28: 192–217
- Matasović, Ranko 2009, *Etymological Dictionary of Proto-Celtic*, Brill, Leiden
- Matasović, Ranko 2012, "Dybo's Law in Proto-Celtic," *Zeitschrift für celtische Philologie* 59: 129–141
- McCone, Kim 1980, "The nasalizing relative clause with object antecedent in the glosses," *Ériu* 31: 10–27
- McCone, Kim 1985, "The absolute and conjunct verbal inflection in Old Irish," in: Schlerath, B. (ed.), *Grammatische Kategorien: Funktion und Geschichte*, Reichert, Wiesbaden: 261–270

- McCone, Kim 1986, "From Indo-European to Old Irish: Conservation and innovation in the verbal system," in: Evans *et al.* 1986: 222–266
- McCone, Kim 1987, *The Early Irish Verb*, An Sagart, Maynooth
- McCone, Kim 1991, *The Indo-European Origins of the Old Irish Nasal Presents, Subjunctives and Futures*, Institut für Sprachwissenschaft der Universität, Innsbruck
- McCone, Kim 1994, "An tSean-Ghaeilge agus a Réamhstair," in: McCone, K., McManus, D., Ó Hainle, C., Williams, N. & Breatnach, L. (eds.), *Stair na Gaeilge in ómós do Pádraig Ó Fiannachta*, St Patrick's College, Maynooth: 61–219
- McCone, Kim 1996, *Towards a Relative Chronology of Ancient and Medieval Celtic Sound Change*, Department of Old and Middle Irish, National University of Ireland, Maynooth
- McCone, Kim 2005, *L'irlandese antico e la sua preistoria* (transl. Elisa Roma), Edizioni dell'Orso, Alessandria
- McCone, Kim 2006, *The Origins and Development of the Insular Celtic Verbal Complex*, Department of Old Irish, National University of Ireland, Maynooth
- McCone, Kim 2015, "Unstressed vowels and consonant quality in Old Irish: U or Non-U," in: Breatnach, L., Ó hUiginn, Ruairí, McManus, Damian, & Simms, Katharine (eds.), *Proceedings, XIV International Congress of Celtic Studies*, Dublin Institute for Advanced Studies, Dublin: 109–135
- McManus, Damian 1991, *A Guide to Ogam*, An Sagart, Maynooth
- Meid, Wolfgang 1968, "Indo-European and Celtic," *Scottish Studies* 12: 45–56
- Meid, Wolfgang 1986, "The Celtic languages," in: Schmidt 1986a: 116–122
- Meid, Wolfgang 1989, *Archäologie und Sprachwissenschaft*, Institut für Sprachwissenschaft der Universität, Innsbruck
- Meid, Wolfgang 1994a, *Celtiberian Inscriptions*, Archaeolingua, Budapest
- Meid, Wolfgang 1994b, "Die 'große' Felsinschrift von Peñalba de Villastar," in: Bielmeier, R. & Stempel, R. (eds.), *Indogermanica et Caucasica: Festschrift für Karl Horst Schmidt*, de Gruyter, Berlin: 385–394
- Mullen, Alex 2007, "Evidence for written Celtic from Roman Britain: a linguistic analysis of *Tabella Sulis* 14 and 18," *Studia Celtica* 41: 31–45
- Ó hUiginn, Ruairí 1986, "The Old Irish nasalizing relative clause," *Ériu* 37: 33–87
- Pokorny, Julius 1959, *Indogermanisches etymologisches Wörterbuch*, Francke, Bern
- Polomé, Edgar C. 1983, "Celto-Germanic isoglosses (revisited)," *Journal of Indo-European Studies* 11: 281–298
- Powell, T. G. E. 1980, *The Celts*, 2nd edn, Thames & Hudson, London
- Prosdocimi, Aldo L. 1989, "L'iscrizione gallica del Larzac e la flessione dei temi in -a, -i, -ja," *Indogermanische Forschungen* 94: 190–206
- Prósper, Blanca María 2002, "La gran inscripción rupestre celtibérica de Peñalba de Villastar: Una nueva interpretación," *Palaeohispanica* 2: 213–226, 421
- Prósper, Blanca María 2008, *El Bronce celtibérico de Botorrita*, Fabrizio Serra, Pisa–Rome
- Raybould, Marilynne E. & Sims-Williams, Patrick 2009, *Introduction and Supplement to the Corpus of Latin Inscriptions of the Roman Empire Containing Celtic Personal Names*, CMCS Publications, Aberystwyth
- Renfrew, Colin 1987, *Archaeology and Language*, J. Cape, London
- Repanšek, Luka 2013, "Two notes on Old Celtic morphology," *Acta Linguistica Petropolitana* 10: 239–254
- Ringe, Donald A. 1988, "Laryngeal isoglosses in the western Indo-European languages," in: *Bammesberger* 1988: 415–441

- Rodway, Simon 2013, *Dating Medieval Welsh Literature: Evidence from the Verbal System*, CMCS Publications, Aberystwyth
- Russell, Paul 1990, *Celtic Word-Formation: The Velar Suffixes*, Dublin Institute for Advanced Studies, Dublin
- Russell, Paul 1995, *An Introduction to the Celtic Languages*, Longman, Harlow
- Schmidt, Karl Horst 1974, "Zur Vorgeschichte des keltischen Kasussystems," *Bulletin of the Board of Celtic Studies* 25: 402–407
- Schmidt, Karl Horst 1979, "On the Celtic languages of Continental Europe," *Bulletin of the Board of Celtic Studies* 28: 189–205
- Schmidt, Karl Horst 1984, "Keltisch und Germanisch," in: Untermann, J. & Brogyanyi, B. (eds.), *Das Germanische und die Rekonstruktion der indogermanischen Grundsprache*, Benjamins, Amsterdam: 113–153
- Schmidt, Karl Horst 1985, "Keltisch, Baltisch und Slavisch," in: Melena, José L. (ed.), *Symbolae Ludovico Mitxelena septuagenario oblatae*, vol. I, Instituto de Ciencias de la Antigüedad/Aintzinate-Zientzien Institutua, Universidad del País Vasco Euskal Erriko Unibertsitatea, Vitoria: 23–29
- Schmidt, Karl Horst (ed.) 1986a, *Geschichte und Kultur der Kelten*, C. Winter, Heidelberg
- Schmidt, Karl Horst 1986b, "The Celtic languages in their European context," in: *Evans et al.* 1986: 199–221
- Schmidt, Karl Horst 1988, "On the reconstruction of Proto-Celtic," in: *MacLennan* 1988: 231–248
- Schmidt, Karl Horst 1990, "On the prehistory of aspect and tense in Old Irish," *Celtica* 21: 593–603
- Schmidt, Karl Horst 1991, "Latin and Celtic: genetic relationship and areal contacts," *Bulletin of the Board of Celtic Studies* 38: 1–19
- Schmidt, Karl Horst 1995, Review of *Études celtiques* 28–29, *Zeitschrift für celtische Philologie* 47: 248–254
- Schrijver, Peter 1992, "The chronology of the loss of post-posttonic vowels between identical consonants and the origin of the Celtic first person singular imperfect," *Münchener Studien zur Sprachwissenschaft* 53: 179–196
- Schrijver, Peter 1993, "On the development of vowels before tautosyllabic nasals in primitive Irish," *Ériu* 44: 33–52
- Schrijver, Peter 1995, *Studies in British Celtic Historical Phonology*, Rodopi, Amsterdam
- Schrijver, Peter 1997, *Studies in the History of Celtic Pronouns and Particles*, Department of Old Irish, National University of Ireland, Maynooth
- Schrijver, Peter 2007, "Notes on British Celtic comparatives and their syntax," in: Nussbaum, Alan J. (ed.), *Verba Docenti: Studies in Historical and Indo-European Linguistics presented to Jay H. Jasanoff*, Beech Stave Press, Ann Arbor, NY: 307–319
- Schrijver, Peter 2014, *Language Contact and the Origins of the Germanic Languages*, Routledge, New York–London
- Schrijver, Peter 2015 "Pruners and trainers of the Celtic family tree: the rise and development of Celtic in the light of language contact," in: Breatnach, L., Ó hUiginn, Ruairí, McManus, Damian, & Simms, Katharine (eds.), *Proceedings, XIV International Congress of Celtic Studies*, Dublin Institute for Advanced Studies, Dublin: 191–219
- Schumacher, Stefan 2000, *The Historical Morphology of the Welsh Verbal Noun*, Department of Old Irish, National University of Ireland, Maynooth
- Schumacher, Stefan 2004, *Die keltischen Primärverben*, Innsbrucker Beiträge zur Sprachwissenschaft, Innsbruck

- Schumacher, Stefan 2009, "Lexical and structural language-contact phenomena along the Germano-Celtic transition zone," in: *Kelten am Rhein: Akten des dreizehnten Internationalen Keltologiekongresses*, Volume II, Philipp von Zabern, Mainz am Rhein: 247–266
- Sims-Williams, Patrick 1981, "The development of the Indo-European voiced labiovelars in Celtic," *Bulletin of the Board of Celtic Studies* 29: 201–229, 690
- Sims-Williams, Patrick 1984, "The double system of verbal inflexion in Old Irish," *Transactions of the Philological Society*: 138–201
- Sims-Williams, Patrick 1990, "Dating the transition to neo-Brittonic: Phonology and history, 400–600," in: Bammesberger, A. & Wollmann, A. (eds.), *Britain 400–600: Language and History*, C. Winter, Heidelberg: 217–261
- Sims-Williams, Patrick 1991, "The emergence of Old Welsh, Cornish and Breton orthography, 600–800: The evidence of archaic Old Welsh," *Bulletin of the Board of Celtic Studies* 38: 20–86
- Sims-Williams, Patrick 1995, "Indo-European \*g<sup>wh</sup> in Celtic, 1894–1994," in: Eska, Gruffydd & Jacobs 1995: 196–218
- Sims-Williams, Patrick 1998, "Celtomania and Celtoscepticism," *Cambrian Medieval Celtic Studies* 36: 1–35
- Sims-Williams, Patrick 2003, *The Celtic Inscriptions of Britain: Phonology and Chronology, c. 400–1200*, Blackwell, Oxford
- Sims-Williams, Patrick 2007a, *Ancient Celtic Place-Names in Europe and Asia Minor*, Blackwell, Oxford
- Sims-Williams, Patrick 2007b, *Studies on Celtic Languages before the Year 1000*, Aberystwyth, CMCS Publications
- Sims-Williams, Patrick 2011, "Celto-Etruscan Speculations," in: Luján, Eugenio R. & García Alonso, Juan Luis (eds.), *A Greek Man in the Iberian Street: Papers in Linguistics and Epigraphy in Honour of Javier de Hoz*, Innsbrucker Beiträge zur Sprachwissenschaft, Innsbruck: 275–284
- Sims-Williams, Patrick 2012a, "Bronze- and Iron-Age Celtic-speakers: what don't we know, what can't we know, and what could we know? Language, genetics and archaeology in the twenty-first century," *Antiquaries Journal* 92: 427–449
- Sims-Williams, Patrick 2012b, "Celtic Personal Names," in: Meißner, Torsten (ed.), *Personal Names in the Western Roman World*, Curach Bhán, Berlin: 151–166
- Sims-Williams, Patrick 2013, "The Celtic Composition Vowels -o- and -io-," in: *García Alonso* 2013: 37–50
- Sims-Williams, Patrick 2015, "The four types of Welsh *yn*," *Transactions of the Philological Society* 113: 286–304
- Sims-Williams, Patrick 2016, "The location of the Celts according to Hecataeus, Herodotus, and other Greek writers," *Étude celtiques* 42 (in press)
- Stalmaszczyk, Piotr & Witzak, Krzysztof T. 1995, "Celtic-Slavic language connections: new evidence for Celtic lexical influence upon Proto-Slavic," *Linguistica Baltica* 4: 225–232
- Stifter, David 2008, Review of Lambert & Pinault 2007, in: *Keltische Forschungen* 3: 267–291
- Stifter, David 2012, "Lenition of *s* in Gaulish?," in: Whitehead, B.N., Olander, T., Olsen, B.A. & Rasmussen, J.E. (eds.), *The Sound of Indo-European*, Museum Tusulanum Press, Copenhagen: 523–544
- Stüber, Karin 1998, *The Historical Morphology of n-Stems in Celtic*, Department of Old Irish, National University of Ireland, Maynooth

- Ternes, Elmar (ed.) 2011, *Brythonic Celtic – Britannisches Keltisch*, Hempen Verlag, Bremen
- Thurneysen, Rudolf 1946, *A Grammar of Old Irish*, Dublin Institute for Advanced Studies, Dublin
- Tovar, Antonio 1986, “The Celts in the Iberian peninsula: Archaeology, history, language,” in: *Schmidt* 1986a: 68–101
- Uhlich, Jürgen 1993, *Die Morphologie der komponierten Personennamen des Altirischen*, M. Wehle, Witterschlick–Bonn
- Uhlich, Jürgen 2007, “More on the linguistic classification of Lepontic,” in: *Lambert & Pinault* 2007: 373–411
- Vendryes, J. 1911, “La place du verbe en celtique,” *Mémoires de la Société de Linguistique de Paris* 17: 337–351
- Vendryes, J., Bachellery, E. & Lambert, P.-Y. 1959–, *Lexique étymologique de l’irlandais ancien*, Dublin Institute for Advanced Studies, Dublin
- Villar, Francisco 1991, “Le locatif celtibérique et le caractère tardif de la langue celtique dans l’inscription de Peñalba de Villastar,” *Zeitschrift für celtische Philologie* 44: 56–66
- Villar, Francisco 1995, *A New Interpretation of Celtiberian Grammar*, Institut für Sprachwissenschaft der Universität, Innsbruck
- Viret, Jérémie; Lambert, Pierre-Yves; Stüber, Karin; Stifter, David & Repanšek, Luka 2013, “Le plomb de Chartres,” *Études celtiques* 39: 123–192
- Watkins, Calvert 1962, *Indo-European Origins of the Celtic Verb*, Volume I, Dublin Institute for Advanced Studies, Dublin
- Watkins, Calvert 1963, “Preliminaries to a historical and comparative analysis of the syntax of the Old Irish verb,” *Celtica* 6: 1–49
- Watkins, Calvert 1966, “Italo-Celtic revisited,” in: Puhvel, J. & Birnbaum, H. (eds.), *Ancient Indo-European Dialects*, University of California Press, Berkeley–Los Angeles: 29–50
- Weiss, Michael 2012, “Italo-Celtica: linguistic and cultural points of contact between Italic and Celtic,” in: Jamison, Stephanie W., Melchert, H. Craig & Vine, Brent (eds.), *Proceedings of the 23rd Annual UCLA Indo-European Conference*, Hempen, Bremen: 151–173
- Willis, David W.E. 1998, *Syntactic Change in Welsh: A Study of the Loss of Verb-Second*, Clarendon Press, Oxford
- Zair, Nicholas 2012, *The Reflexes of the Proto-Indo-European Laryngeals in Celtic*, Brill, Leiden
- Ziegler, Sabine 1994, *Die Sprache der altirischen Ogam-Inschriften*, Vandenhoeck & Ruprecht, Göttingen
- Zimmer, Stefan 2000, *Studies in Welsh Word-Formation*, Dublin Institute for Advanced Studies, Dublin

# GERMANIC

*Joshua Bousquette and Joseph Salmons*

We survey Germanic as a distinct branch of Indo-European, tracing developments that differentiate Germanic from IE up to attested Germanic varieties. We cover structural patterns shared across parts of the family and those that distinguish the sisters from one another. Beyond familiar examples (umlaut, case loss, verb second), we treat areas of recent progress, such as laryngeal phonetics and phonology, foot structure, definiteness and complementizers. We stress how these patterns cut across grammatical modules, connecting phonology, morphology and syntax. That discussion is embedded in the available social and cultural context, e.g. language contact.

## INTRODUCTION

Our earliest evidence for Germanic languages is from names and words recorded in classical-language sources, including Tacitus, Caesar and Pliny. Most discussed are tribal, personal and place names, but they also include cultural vocabulary, some borrowed into Latin, e.g. *framea* ‘spear’, *glēsum* ‘amber’ (cf. *glass*) and possibly cognates of *soap* ‘red hair dye’ (Green 1998: 185–188, Kluge 2011).

Germanic languages are also attested in runic inscriptions written by speakers in the first centuries of the Common Era (Krause 1971, Antonsen 1975, Schulte 2006). These are not limited to Scandinavia (as with the Negau helmet, an early inscription, found in present-day Slovenia) and are identified with various Germanic languages, based on where they were found (problematic for inscriptions on portable objects like spearheads and brooches), orthography and structure. Page (2001) illustrates these difficulties for Frisian runic inscriptions, where arguments for identifying inscriptions as Frisian are tentative and limited to a few items; e.g. the evidence for Frisian provenance of the apparent personal name *skanomodu* assumes a connection to Gmc \*skaun- ‘beautiful’, showing distinctively Frisian monophthongization.

The first longer text preserved in a Germanic language is part of a Bible translation, attributed to Wulfila (Ulfilas) in the late 4th century, written in his Greek-based script. This is our major source for Gothic, in a 6th century Italian manuscript. Texts in the Latin alphabet begin after 500 CE, summarized here, with rough dates for composition and surviving manuscripts:

### (1) Early longer texts in selected Germanic languages

Gothic	late 4th c., manuscript from the 6th c.
Old English	mid–late 7th c., manuscript from the 8th c.
Old High German	mid 8th c., manuscript from the late 8th c.
Old Saxon	ca. 830, major manuscript from the mid 11th c.
Old Norse	mid 11th c.
Old Frisian	late 11th c., manuscript from the late 13th c.



The family is traditionally divided into three branches:

- East Germanic: Gothic and fragments from other languages. The last known survivor was “Crimean Gothic”, recorded in 1562 (Stearns 1978).
- North Germanic: Runic and then Old Norse, represented by the modern Nordic languages.
- West Germanic, including English, German, Dutch, Yiddish, Frisian, etc.

These branches show their characteristics in the earliest attestations; see Robinson (1992: ch. 10). He eschews sub-branching within this tree, yielding the picture on the left; another view is that on the right (Salmons 2012: 84), a three-way division where North and West Germanic are closer to each other than to East Germanic.



FIGURE 8.1 TWO MODELS OF GERMANIC SUBGROUPING

Sociohistorical evidence supports the model on the right, with the relatively early departure of East Germanic from southern Scandinavia (Heather 2010), leaving a dialect continuum across North and West Germanic.

One gauge for subgrouping is shared innovations that are unlikely to be independent innovations, but these prove problematic. In sound change, glide hardening – the development of stops from glides (*Verschärfung*, Holtzmann’s Law), shared between Gothic and North Germanic – has long seemed exotic. In both languages, reconstructed geminate glides (\**jj*, \**ww*) appear as geminate stops, so that OHG *triuwa* (cf. *troth*) corresponds to ONor. *tryggva* and Goth. *triggwa*. Yet, while similar reflexes do appear, the outcomes do not always match up: OHG *zweiio* ‘of two’ corresponds to coronals in Goth. *twaddje* and velars in Norse *veggja*. And Faroese (but not its cousin Icelandic) underwent a later sharpening (*skerping*), where apparently geminate *-jj-* and *-ww-* hardened eventually into fortis geminates, palatal and velar respectively (data and analysis following Árnason 2011: 31–33, who uses Old Icelandic as a starting point):

(2) Faroese *skerping* (adapted from Árnason 2011: 31)

Old Icelandic	Intermediate stage	Modern Faroese	
eyjar [øyjar]	[dʒ:]	oyggjar [ɔ(i)tʃ:a]	‘islands’
róa [row.wa]?	[gʷ:]	rógva [ɹɛkʷ:a]	‘to row’

Worse, hardening is hardly foreign to West Germanic, though in different environments. Schirmunski (2010: 429) shows southern Swabian hardening of labiovelar glides to *b*, cf. *aib̥ar* ‘your (pl.)’, MHG *iower*. In Low German, palatal glides can harden to a spirant, e.g. Soest [ɛʏa] *eier* ‘eggs’ < OS *eiero* < PGmc \**ajj-* (Hall 2014). Modern North Germanic, moreover, seems to show similar tendencies. Riad (2014: 21, 29) discusses the rise of “damped” or “buzzing” realizations of [i:] and [y:] as [i:ʔ] and [y:ʔ] and of a [β] off-glide

after long high back vowels: [fiːn] for *fin* ‘fine’ and [buːβk] for *buk* ‘book’. Here, we have glide hardening without motivation from syllable structure, since these are additional coda material rather than new onsets. High vowels and glides harden into stops around the world; e.g. earlier Inscriptional Burmese *-iy* and *-uw* correspond to *-it* and *-uk* in Maru (Mortensen 2012).

To take another example, rhotacism, the development of /r/ sounds from earlier \*z, is shared by North and West Germanic: Gmc \*deuza ‘animal’, Goth. *dīus*, OHG *tior* and ONor. *dýr*. Here, too, the change is far from unique, as in Lat. *flōs* ~ *flōrem* plus later, independent occurrences of rhotacism in Romance, Germanic and other IE languages (Catford 2001), undercutting its value as a diagnostic.

In morphology, innovations are often built from inherited material in ways that could be independent, like class iv weak verbs (*-nan*) in Goth. *gadauþnan* ‘die’ and Norse *vakna* ‘awaken’, absent in West Germanic. Conversely, the innovative *-st* 2 sg. verbal inflection is characteristic of West Germanic only, attested first in Old High German in high-frequency verbs in main clauses, and later spreading to other tenses, moods and syntactic positions (Somers 2011). Independent developments in single branches may further cloud relationships. Gothic exhibits a synthetic middle passive, e.g. (*ik*) *haitada* ‘I am called’, where *-da* may be reflexive, consistent with reconstructed IE (p. 162). This is already in competition with reflexive verbal constructions using *sik*, *sis*, or *seina* ‘self’, a process that occurred in Greek (Ferraresi 2005: 106–109); synthetic passives are absent in later Germanic. (The Scandinavian middle is a later development, a reanalysis of the reflexive pronoun *sik* as verbal *-sk*; Faarlund 2004: 126–127, also Emonds & Faarlund 2014.)

If a unified Germanic once existed, it was perhaps ca. 500 BCE. Early Runic shows surprising uniformity over time and space, possibly due to conservative writing traditions and a limited early corpus. The evidence shows differentiation across branches from the beginning; some West Germanic varieties show nasal loss before coda fricatives (/f, θ, x/) with compensatory lengthening of the preceding vowel, while others do not: Eng. *five*, *mouth* vs. Germ. *fünf*, *Mund*, from Gmc \*femfa, \*munθa; a number of forms show variable reflexes, cf. MHG *sunt* ~ *sūd*.

While a tree-like split into three distinct branches is well supported (Nielsen 2000, Heather 2010), intra-Germanic contact was nevertheless pervasive. “North Sea Germanic” (Ingvaenic) areas still form dialect continua, crossing the Dutch-German linguistic border. And rich patchwork distributions of features occur across mainland Scandinavia. This has blurred the lines of a Stammbaum with chronic wave-like diffusion.

More dramatically, English was initially formed by a mixture of Germanic varieties during the invasion of the British Isles, shaped by the shift of Celtic (“British”) speakers to English, then the North Germanic invasions during the Danelaw and eventually the Normans. What we now call English came into existence thanks to sociohistorical settings involving high numbers of bilingual and bidialectal speakers and second language and second dialect speakers, dialect contact (koineization or “new dialect formation”; Kerswill 2002) plus language contact (Trudgill 2010). This has led to the proposal that it has become North Germanic (Faarlund 2012) or a “fourth branch of Germanic” (Forster et al. 2006). These claims have garnered press attention, but the mainstream view remains that English is West Germanic with complex effects of historical contact (McWhorter 2002, Trudgill 2010).

Some characteristics of Germanic are often tied to the sociohistorical scenarios just sketched. Consider two examples. Substrate effects are likely in lexical material from pre-IE-speaking populations, and many terms reflect domains where vocabulary survivals

from earlier languages seem particularly plausible (flora, fauna, cultural vocabulary); Polomé (1989) and Hamp (1990) argue for substrate status when words show aberrant phonological and/or morphological patterns vis-à-vis inherited vocabulary. Hamp lays out, for instance, the difficulties of the ‘apple’ word, showing reflexes of the extremely rare \*b phoneme (also Salmons 2004, 2015). Let us turn to the more secure lexical connections with attested languages, Finnic and other IE groups.

Germanic loans in Baltic Finnic show great antiquity and have been explored for what they show about Proto-Germanic (Kylstra 1991–, Koivulehto 2001, both in some respects controversial). Finnish forms like *kuningas* ‘king’ (Germ. *König*) and *ringas* ‘ring’ retain nominal stem vowels plus endings lost in Germanic. Some words suggest pre-Germanic forms, like Finn. *kana* ‘hen’ with an unshifted \*k, rather than Gmc \*h.

Borrowings from Germanic into Romance similarly allow us to date borrowings and identify the nature of contact. For example, early borrowings with word-initial labial glides from Germanic into Romance predate glides hardening: Old Frankish \*wra $\dot{c}$ jo ‘exile’ (OHG *reccheo*) becomes Old French *garçun* ‘servant’; Gmc \*werre ‘war’ (Goth. *warjan* ‘defend’, cf. Eng. *warrior*, *warden*) becomes Old French *guerrer* (cf. Eng. *guerilla*, *guard*). These cognates in English (post 1066) attest to borrowing from Germanic into Romance, then back into Germanic.

Germanic-Celtic contact predated the Germanic breakup, as all branches share some of the same cognates with Celtic. Borrowed from Celtic are some legal terms, including ancestors of words like *oath* (Lehmann 1986). Goth. *andbahts* ‘service, office’ (Germ. *Amt*) is distinguishable as a borrowing by the /a/ for IE syllabic \* $\text{ṛ}$  (IE \* $\text{-ṛ}^{\text{h}}$ ), which has anaptyctic /u/ in Germanic (OHG *umbi* ‘around’, but Gr.  $\acute{\alpha}\mu\phi\iota$ ). Another likely early borrowing is OIr. *dún* ‘fort, walled hillock’ (W *dinas*, ‘city’): OHG *zūn* ‘fence’ (Germ. *Zaun* ‘fence’), OEng./OS/ONor. *tūn* ‘village’ (Dutch *tuin* ‘garden, yard’). *Zaun* reflects the second sound shift, so predates that. While some accept such terms as Celtic loans (Orel 2003, Kroonen 2013), others are skeptical, e.g. Kluge 2011, where *Zaun* is of “uncertain origin” but presumably connected to Celtic. A third is more complex, Goth. *eisarn-* ‘iron’, OEng. *isarn/isærn*, OHG *isarn*, MDutch *iser* ‘iron’ (OIr. *iarn*). (The iron age came relatively late to Northern Europe, perhaps 400 BCE, and some connection to Celtic is likely; Kluge (2011) suggests that Celtic and Germanic borrowed the term from a third language, while Lehmann (1986) notes possible multiple borrowings from Celtic.)

Contact often also correlates with morphological reductions, e.g. case loss (O’Neil 1978). A standard narrative about Germanic is framed around such “simplifications”. As argued below, that view must be balanced against INCREASING complexity, including in the segmental inventory (larger vowel systems) and new inflection. Simplification is a frequent result of contact, but hardly inevitable (see below).

Germanic’s position within IE has been more controversial (see Polomé 1972). Some (e.g. Ringe 2006: 5–6) assign Germanic to a “central” group including Balto-Slavic, Indo-Iranian, Armenian and Greek. Vocabulary is shared in complex ways with north-western neighbors (Meillet 1967), perhaps reflecting later contacts rather than earlier genetic unity. Contact took place with pre-IE languages, other IE branches and non-IE languages, especially Finnic (Roberge 2010, Mailhammer & Vennemann 2015), muddying the waters between tree- and wave-like patterns. Like other IE languages, Germanic was likely forged from contact between Indo-Europeans and indigenous populations. Jutland (Denmark), southernmost Norway and Sweden, and the Baltic and North Sea coasts are likely the earliest area where Germanic was spoken, with ongoing contact with other (post-)IE groups and non-IE groups.

## PHONOLOGY

### Consonants

Germanic stops and fricatives differ systematically from IE, other segments less so. The IE laryngeals, on mainstream views, were lost early and have little relevance to Germanic (but see Müller 2007). Traces of laryngeal phonology are visible in the ablaut system (Ringe 2006: 80):

(3) PIE root-ablaut alternations		pre-PGmc root-ablaut alternations
a ~ Ø	>	a ~ Ø
e ~ Ø ~ o	>	e ~ Ø ~ o
h <sub>1</sub> e ~ h <sub>1</sub> ~ h <sub>1</sub> o	>	e ~ Ø ~ o
h <sub>2</sub> e ~ h <sub>2</sub> ~ h <sub>2</sub> o	>	a ~ Ø ~ o
h <sub>3</sub> e ~ h <sub>3</sub> ~ h <sub>3</sub> o	>	o ~ Ø ~ o
eh <sub>1</sub> ~ h <sub>1</sub> ~ oh <sub>1</sub>	>	ē ~ a ~ ō
eh <sub>2</sub> ~ h <sub>2</sub> ~ oh <sub>2</sub>	>	ā ~ a ~ ō
eh <sub>3</sub> ~ h <sub>3</sub> ~ oh <sub>3</sub>	>	ō ~ a ~ ō

Similarly, whatever the status of IE palatal versus velar stops, both surface in Germanic as velars (Sihler 1995: 151–168):

(4) IE	English	
*yugom-	<i>yoke</i>	‘yoke’
*ǵenh <sub>1</sub> -	<i>kin</i>	‘beget’
*ǵneh <sub>3</sub> -	<i>know</i>	‘know’

IE syllabic resonants – liquids and nasals – resolve into \*u plus resonant (Kluge 2011):

(5) pre-Germanic	Germanic	
*wǝrm-	*wurma-	‘worm’
*wǝlk <sup>w</sup> o-	*wulfa-	‘wolf’
*ǵndher-	*under-	‘under’
*smH-	*sumera	‘summer’

This leaves this pre-Germanic system:

TABLE 8.1 PRE-GERMANIC CONSONANTS

	<i>Labial</i>	<i>Coronal</i>	<i>Palatal</i>	<i>Velar</i>	<i>Labiovelar</i>
<b>stops</b>					
voiceless	p	t		k	k <sup>w</sup>
voiced	b	d		g	g <sup>w</sup>
murmured	b <sup>h</sup>	d <sup>h</sup>		g <sup>h</sup>	g <sup>wh</sup>
<b>fricative</b>		s			
<b>sonorants</b>					
liquid		l, r			
nasal	m	n			
glide	w		j		

Traditional IE obstruents differ from, yet correspond systematically to, Germanic:

TABLE 8.2 IE AND GERMANIC STOPS

	IE	Germanic
<b>voiceless</b>		
labial	*p	*f
coronal	*t	*θ
velar	*k	*x
labiovelar	*k <sup>w</sup>	*x <sup>w</sup> > *h <sup>w</sup>
<b>voiced</b>		
labial	*b	*p
coronal	*d	*t
velar	*g	*k
labiovelar	*g <sup>w</sup>	*k <sup>w</sup>
<b>voiced aspirate</b>		
labial	*b <sup>h</sup>	*b / *β
coronal	*d <sup>h</sup>	*d / *ð
velar	*g <sup>h</sup>	*g / *ɣ
labiovelar	*g <sup>hw</sup>	*g <sup>w</sup> / *ɣ <sup>w</sup>

In these changes (“Grimm’s Law”, Germanic Consonant Shift, First Sound Shift), one change, \*p > \*f, reflects a change in manner of articulation, from stop to fricative, while others involve laryngeal phonology (meaning “states of the glottis”, i.e. phenomena like voicing and aspiration, not IE laryngeals).

While the shift represents perhaps the most discussed sound change, discussion is often unclear about the precise laryngeal characterization of IE and Germanic stops and fricatives, especially with regard to the division of labor between phonology and phonetics. Crosslinguistically, laryngeal contrasts can be captured with three features, [voice], [spread glottis] and [constricted glottis], plus laryngeally unmarked or unspecified consonants (Iverson & Salmons 1995, Avery & Idsardi 2001). In many languages, these contrasts are enhanced by additional phonetic cues, particularly for phonation (Henton et al. 1992). The traditional IE obstruent system would have had unmarked stops (\*p, \*t, \*k, etc.), voiced stops (\*b, \*d, \*g, etc.) and stops marked by both [voice] and [spread glottis] (\*b<sup>h</sup>, \*d<sup>h</sup>, \*g<sup>h</sup>, etc.). Earlier voiceless stops become voiceless fricatives (“spirantization”), perhaps triggered by aspiration, as an enhancement. (A shift to a [spread glottis] system may have preceded these changes.) IE voiced stops devoice, a process found in many languages, perhaps motivated by the challenge of maintaining voicing – which requires considerable air flow across the glottis, while air flow is blocked in stops. The third are traditional “voiced aspirates”; modern descriptions of Indic treat them as “murmured” or “breathy”, with a slightly open glottis with high air flow; these deaspirate in Germanic. Controversy surrounds whether these become voiced stops or fricatives. For instance, a form like \*b<sup>hend</sup> becomes Gmc \*bind-, with voiced stops initially and after a nasal, but postvocally we often find fricatives, e.g. ONor. *orð* ‘word’ from \*wrd<sup>h</sup>o-. We assume underspecification for [continuant], similar to contextual stop-fricative alternations in Spanish.

These changes – spirantization, devoicing, deaspiration – appear to be connected by encroachment of new sounds into the space of old ones or the filling of gaps created by earlier changes, “push” or “pull” chains respectively. On the former, if \*d<sup>h</sup> deaspirated first, \*d<sup>her</sup>- would have risked merger with \*der-, triggering devoicing of \*der- to \*ter-. On the

latter, if \*pel- become \*fel-, the lack of voiceless stops might allow devoicing to fill this gap. This pull is sometimes motivated in terms of voiceless stops as “unmarked”, lacking laryngeal specification. (Wrinkles include labiovelars; see Salmons & Smith 2005.)

Germanic contrasts [spread glottis] or “aspirated” \*p, \*t, \*k, etc. and unmarked \*b, \*d, \*g, etc., rather than “voiceless” vs. “voiced”, making Germanic phonologically more different from IE than traditional views, with implications for later obstruent changes.<sup>1</sup> (Dutch and Yiddish are [voice] languages, in both languages likely due to contact with Romance and Slavic, respectively.)

A set of Germanic forms shows irregular voicing in fricatives, which once constituted a challenge to claims about the regularity of sound change:

(6) Verner’s Law

Skr. <i>bhrátar-</i>	Gmc *brōþer	‘brother’
Skr. <i>mātár-</i>	Gmc *mōðer	‘mother’

Verner (1875) famously established that this voicing was actually regular, conditioned by position of the IE “mobile” accent (§2.3), so that (expected) voiceless fricatives appeared where Sanskrit has a preceding accent and voiced when the accent did not precede. Germanic provides evidence for the position of the IE accent in verbal paradigms, reflected in a few modern alternations like Eng. *was* ~ *were* (/r/ from \*z). Unmarked stops (Germ., Eng., Dan. <b, d, g>) are, as noted, treated as unspecified. They often lack phonetic voicing in initial and final position (systematically in Icelandic), but in a voice-friendly environment like between vowels, they are phonetically voiced. A similar kind of “passive voicing” (Iverson & Salmons 2003, among others) may have been at work in Verner’s Law: IE accent was likely realized with stiff vocal folds, so that a preceding accent would inhibit fricative voicing.

Another “exception” is equally important, but never played the pivotal role that Verner patterns did concerning regularity: IE voiceless stops did not shift after \*s: \*sper- retains its stop in modern Eng. *spear*, Germ. *Speer*. The modern absence of aspiration – or phonological specification for [spread glottis] – supports connecting spirantization to aspiration.

These developments illustrate continuity and change within Germanic. Complex consonantal “chain shifts” like Grimm’s Law are uncommon in the world’s languages, but occur again within later Germanic. High German is defined traditionally by a set of changes known often as the Second Consonant Shift (“High German Consonant Shift”). Like in Grimm, stops become spirants, though variably by environment and place of articulation, and in certain varieties and context showing a partial shift to affricate rather than fricative. Intervocalic stops are prone to shift but initial or post-consonantally less so, and coronals preferably over velars, as in English (unshifted) and German (shifted):

(7) The Second Consonant Shift

English	German
<i>path</i>	<i>Pfad</i>
<i>help</i>	<i>helfen</i>
<i>town</i>	<i>Zaun</i>
<i>eat</i>	<i>essen</i>
<i>cool</i>	<i>kühl</i>
<i>milk</i>	<i>melken</i>
<i>book</i>	<i>Buch</i>

Modern parallels include affrication of initial /t/ in some varieties of Danish and affrication/spirantization in Liverpool English (e.g. Honeybone 2002).

The Second Sound Shift has a broader chain-like character. The lenis (in traditional terms “voiced”) \*d becomes /t/ except in the north: Eng. *drink* is cognate with Germ. *trinken*. Farther south, \*g becomes /k/, and \*b becomes /p/: Old Bavarian *perec* ‘mountain’, cf. *Berg*. Possibly related is the “stopping” of interdental \*θ across most of Germanic. Excepting variation based on dialect and the environment, the basic progressions of the sound shifts are illustrated in (8).

(8)

Indo-European	Germanic	Old High German
b <sup>h</sup> d <sup>h</sup> g <sup>h</sup>		
b d g	b d g	b d g
p t k	p t k	p t k
	f þ x>h	pf ts kh
		f d h
s	z	s
		r
sp st sk		sp st sk

Other widespread patterns include final laryngeal neutralization, “final devoicing” (Iverson & Salmons 2011). Some patterns of variability, like a wide range of realizations of rhotics across languages and dialects, continue ancient variation (Howell 1991).

Overall, attested changes often parallel those posited for prehistory and early history, with occasional twists on claims about innovation: while English is often tagged as an ill-behaved member of the family, its obstruent system remains surprisingly close to West Germanic.

## Vowels

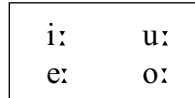
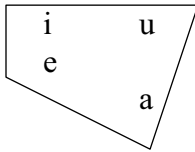
Early Germanic maintains a vowel system close to that of IE after the emergence of distinct long vowels, with a five-vowel system with a length contrast. Prehistoric low-back vowels merged, in different directions by length:

(9) (late) Indo-European >	Germanic	
*g <sup>h</sup> ans	*gans	‘goose’
*orb <sup>h</sup> -	*arb-	‘heir, inheritance’
*b <sup>h</sup> rāter	*brōθer	‘brother’
*plō(w)-	*flō-	‘flow’

Such changes have parallels in IE and are reenacted with the contemporary North American merger of the low-back vowels /a/ and /ɔ/.

A further early change is the emergence of a new long vowel, ē<sup>2</sup>. In West Germanic, old \*ē (“\*ē<sup>1</sup>”, occasionally written \*ǣ) had lowered, and the new vowel, attested in Old High German as *ia*, appears in cognates of *here* and as the stem vowel in class VII strong verbs of the *let* type, etc.

## (10) The Proto-Germanic vowel system

**Short vowels**

- \*widuw- 'widow'
- \*fīmf 'five'
- \*et- 'eat'
- \*nas- 'nose'
- \*wull- 'wool'

**Long vowels**

- \*wīd- 'wide, broad'
- \*sīm- 'string'
- \*dēd- 'deed'
- \*brōþer 'brother'
- \*mūs 'mouse'

Earlier \*ey monophthongizes to \*ī. Aside from that, Germanic also included three "diphthongs"; that the \*a in these participates in the low-back merger may suggest that these may have been vowel-glides rather than single units.

## (11) Proto-Germanic "diphthongs"

aj  
ew  
aw

- \*ajh-/ajg- 'have'
- \*skajd- 'separate'
- \*dewp- 'deep'
- \*kews- 'choose'
- \*rawd- 'red'
- \*hawzij- 'hear'

From there, things become steadily more complicated, driven by patterns of vowel to vowel assimilation. Umlaut can be understood only in the context of height harmony and consonantal conditioning. In Germanic, vowels assimilated partially to the height of following vowels, e.g. /e, o/ > /i, u/ before high vowels, as well as before coda nasals (with parallels in some modern dialects), and Gothic shows a general raising of /e/, though /e/ still appears before /h, h<sup>w</sup>, r/. A parallel lowering before /a/ is also illustrated below:

## (12)

- \*geban > OHG *geban* ~ *gibis* 'give ~ you give', but Goth. *giban*
- \*helpan > OHG *helfan* ~ *hilfis* 'help ~ you help', but Goth. *hilpan*
- \*ed- > ONor. *eta*, OHG *ezzan*, but Goth. *itan* 'eat'
- \*bendan > OHG *bindan* 'tie'
- \*wiraz > OHG *wer* 'man'
- \*juka > OHG *joh* 'yoke'

Within Continental West Germanic "primary" umlaut of /a/ to /e/ before i/j shows raising but with additional fronting. That process was blocked dialectally by coda /h, r, l/.



It preceded the general fronting of all back vowels since these are different structural patterns. Moreover, failure of umlaut is attested in southern German dialects (particularly with geminate velars) and in western Netherlandic (blocking environments or  $\bar{a}$ ) (Buccini 1995, Iverson & Salmons 1996, Howell & Salmons 1997). In other daughters, chronological differentiation is not demonstrable, e.g. Old English. Conditioning also differs across the family: North Germanic umlauts heavy stems (*gestr* ‘guest’ < \**gastiz*) but not light ones (*staðr* ‘place’ < \**staðir*) of the relevant morphological classes (Iverson & Salmons 2012). See §3.2 on the morphologization of umlaut.

Above we discussed chain-like consonant shifts, e.g. with fortis or voiceless stops repeatedly becoming affricates and/or fricatives. Similar changes are pervasive in the vowel system. “Chain shifts” have been extensively described for North and West Germanic (Sievers 1876, Küspert 1988, Wiesinger 1970), and related shifts are underway across North America (Labov 1994, among many others). Consistent trends emerge across these independent changes: raising of long or tense vowels, lowering of short or lax vowels and fronting of back vowels.

### Prosody

IE “accent” was marked by pitch and was mobile in that it could fall on different syllables within the paradigm, like in modern Russian (§2.1). The Germanic “Accent Shift” fixes the accent on the initial syllable, obviously after Verner’s Law. Germanic stress correlates broadly with vowel duration and intensity. Kiparsky & Halle (1977) and Halle (1997) provide phonological accounts of the change, while Salmons (1992) argues that language contact played a role, alongside such language-internal motivations.

The effects of this shift are far-reaching, including ties to innovations in verbal and pronominal agreement. For example, Gothic first and second person dual subject pronouns *wit* and *jut* seem to derive from the stems of plural *weis* and *jus* and an enclitic variant of ‘two’, *twáil/twa/twōs*. Clitics were arguably present in earliest Old Norse (Faarlund 2004: 35, though see Biberauer & Roberts 2008: 105) but disappeared with the development of the Scandinavian lexical accent, ca. 800 (Oftedal 1952, Kristoffersen 2011). This accent counteracted the clause-based prosody of other Germanic languages; Scandinavian lexical items determined prosody, independent of clausal position.

West Germanic reduction of final unstressed syllables introduces paradigmatic ambiguity, since Germanic typically marked information on suffixes. Consider the three weak verb classes well attested in West Germanic and distinguished by final vowels (§3.1, §4).

TABLE 8.3 OLD HIGH GERMAN WEAK VERBS

Verb class	Old High German	Modern German
i	<i>fellen</i> ‘fell’	<i>fällen</i> ‘fell’
ii	<i>dionōn</i> ‘serve’	<i>dienen</i> ‘serve’
iii	<i>habēn</i> ‘have’	<i>haben</i> ‘have’

As unstressed vowels reduced to schwa, these distinctions were lost. Such increased ambiguity within inflectional paradigms, often presumed to be driven by reduction and subsequent loss of morphological marking, correlates with an increase in periphrastic forms, e.g. higher frequency of pronominal subjects (Axel & Weiß 2011) and development of obligatory definite articles in Old High German (§5.5).

Other processes create new affixes and inflectional categories. Typically unstressed elements like pronouns cliticize onto a preceding word, creating preferred foot structures,

and pronominal enclitics can be reinterpreted as inflectional affixes or pronouns (De Vogelaer 2010). The OHG 1 pl. verbal *-emes/-emēs*, as in *habemēs*, alternates with a short form, *habēn*, both meaning ‘we have’. The longer inflection is likely from an enclitic *uuis* ‘we’, with assimilation of the coda nasal /n/ or /m/ to the bilabial glide in the onset of the pronoun, yielding /m/. Axel (2007: 305) and Axel & Weiß (2011) show that the long ending regularly appears without a co-occurring subject pronoun, arguing that the long ending is in fact a reanalyzed pronominal form; subject pronouns are obligatory with the “short” ending (see also Early Modern Nuremberg *-ma* (1 pl.); Fertig 2000: 47). Rhotacized forms regularly appear in all Germanic languages other than Gothic, so the innovative inflection with OHG *s* reflects a prehistoric process.

Similar are innovative 1 pl. pronouns, including Flemish, Zeelandic and Brabantic 1 pl. *me*, derived from enclitic *wij/we* and the preceding nasal /n/ of the verbal inflection, which can appear not only phonetically conditioned but also in topic position (De Vogelaer 2010: 15); and Bavarian 1 pl. *mir/ma* from *wir* (Fuß 2004, 2005), which may similarly appear in topic position. Analogous pronominal innovations in North Germanic similarly derived Old Norwegian 1 pl. *mit/mēr* from *wit/wēr*, prevalent between 1200–1400 CE; and Sw. 2 pl. *ni* from *i* during the 17th century (Haugen 1976: 302, 375). Innovative *dir* also appears in numerous dialects of German (Weise 1907: 201ff.), from reinterpretation of 2 pl. inflectional *-t* as part of pronominal *ihr*.

A final prosodically-driven change is complementizer agreement (C-agr), from the reanalysis of pronominal elements as inflectional affixes. C-agr is possible on a number of different lexical categories heading a subordinate clause (Reis 1985, Kathol 2001). C-agr is attested across multiple non-standard West Germanic varieties as early as the mid 13th century (Goeman 1997). An oft-cited example is from Bavarian (Bayer 1984: 233), where the complementizer *wenn* ‘if’ inflects for person and number, here 2 sg.

- (14) *Wennst kummst*  
if-2SG. PRO come-2SG.  
‘If you come’

Similar to innovative inflectional affixes and pronouns, the encliticization of unstressed pronouns to a stressed lexeme is argued to be a historical prerequisite for complementizer agreement (Weiß 2005, forthcoming). Together with innovations in the verbal and pronominal paradigm, C-agr illustrates prosodically based complexification and is relatively stable in non-standard varieties and language islands (Bousquette 2014).

In another area of prosody, syllable and foot structure, the classic insight into Germanic syllable structure is “Prokosch’s Law” (the “Weight Law”), that stressed syllables must contain two moras.<sup>2</sup> Murray & Vennemann (1983) show a conspiratorial pattern of sound changes in all three branches of Germanic consistent with one syllabification of medial consonants: (1) West Germanic gemination, (2) North Germanic vowel lengthening and (3) Gothic glide strengthening.

These effects all fall out, they argue, from the conflict between Prokosch’s Law (which they call the “Stressed Syllable Law”) and the Syllable Contact Law, namely (1983: 520, cf. Vennemann 1988, § = syllable boundary):

- (15) Syllable Contact Law. The preference for a syllable structure  $A\$B$ , where  $A$  and  $B$  are marginal segments and  $a$  and  $b$  are the Consonantal Strength values of  $A$  and  $B$  respectively, increases with the value of  $a$  minus  $b$ .

For example, Icelandic and Faroese show slightly different patterns of syllabification in  $C\acute{V}CCV$ - words, according to consonantal “strength”. Icelandic allows more word-internal

onset clusters, *p, t, k, s + j, v r*, so *vis\$na* ‘wither’ but *vö\$kvva* ‘to water’; Faroese only tolerates *p, t, k + r, l*, except *tl*, so *sjóg\$vrur* ‘sea’ but *sū\$skri* ‘sugar’ (1983: 522–523).

Similarly, Drescher & Lahiri (1991) propose that a wide range of early Germanic sound patterns can be captured in terms of foot structure. They define the “Germanic foot” as “maximally binary, left-headed, where the head must dominate at least two moras” (1991: 251). The head could include one or two syllables, and the non-head element may contain only one mora. That is, a foot required a head of two moras, which could come from one heavy syllable or from two light ones. In their analysis of Old English High Vowel Deletion (simplifying), high vowels delete only in the weak branch of a foot, but those in the head are maintained: CV:C or CVCCV, in OEng. *go\$ð* < \**go\$ðu* ‘good’ or *word* < \**wordu* ‘word’, for example) versus CVCV, in OEng. *loftu* ‘praises’.

In prosody, Germanic and its daughter languages historically also show patterns where prosodically weak, phonetically reduced elements affix to adjacent syllables (as in the cliticization examples above). They thereafter could be incorporated into a “prosodic word” (pword), allowing processes otherwise confined to word boundaries, as noted in Somers Wicka’s (2009: 8) treatment of the Otfrid’s OHG *Evangelienbuch*. Hanna (2009) revisits Jespersen’s Cycle, also in the *Evangelienbuch*, as well as in the MHG *Nibelungenlied*, arguing that “secondary stress [as] the outcome of an interaction between phonology and syntax . . . is sensitive to rhythmic context, to syllable structure, and to vowel height” (2009: 201). Characteristic of Germanic even from early attestations is the phonology’s sensitivity not only to foot structure but also to syntactic environment.

Smith (2004) and Smith & Ussishkin (2014) frame foot and other prosodic structures in terms of templates, allowing us to capture conspiracy-like effects here, just as Murray and Vennemann do for the syllable level.

The traditional view of Germanic prosody is that the accent shift triggered weakening of final syllables leading to the dramatic reductions, “simplifications”, in inflectional morphology, discussed just below. We have also illustrated how reduction can have the opposite effect, increased complexity via cliticization and grammaticalization. More broadly, a simple “phonology first” view of such changes has been questioned within Germanic and beyond recently by Enger (2013) and Menz (2010). Enger (2013: 18) concludes that sound change alone cannot account for reductions in the history of Norwegian, arguing for a driving role for an autonomous morphology.

## MORPHOLOGY

### Inflectional morphology

Germanic strong verbs exploit ablaut for tense. Traditionally divided into 7 classes, the first 5 are inherited from IE and involve a present tense *e*-grade, which alternates with an *a*-grade in the preterit singular, and zero-grade in the preterit plural and past participle.

TABLE 8.4 GERMANIC STRONG VERBS, CLASSES I–V

	Template	Present <i>e</i> -grade	Preterit singular <i>a</i> -grade	Preterit plural Ø-grade	Past participle Ø-grade	Meaning
I	CejC	<i>dreiban</i> <i>reidan</i>	<i>draib</i> <i>raid</i>	<i>dribum</i> <i>ridum</i>	<i>dribans</i> <i>ridans</i>	‘drive’ ‘ride’
II	CewC	<i>leugan</i> <i>sleupan</i>	<i>laug</i> <i>slaup</i>	<i>lugum</i> <i>slupum</i>	<i>lugans</i> <i>slupans</i>	‘lie’ ‘slip’

III	CeLC	helpan	halp	hulpum	hulpans	‘help’
	CeNC	drinkan	drank	drunkum	drunkans	‘drink’
IV	CeL	stelan	stal	stēlum	stulans	‘steal’
	CeN	k <sup>w</sup> eman	k <sup>w</sup> am	k <sup>w</sup> ēmum	k <sup>w</sup> umans	‘come’
V	CeC	geban	gab	gēbum	gebans	‘give’
		wegan	wag	wēgum	wegans	‘move’

This so-called *e*-group is built around the CV templates shown in Table 8.4 (second column). In the absence of a preterit plural root vowel, glides in class I and II develop respectively into a short /i/ and short /u/; in class III the liquids and nasals epenthesize /u/ (§1). The long /e/ in the zero-grade of class V and in the preterit plural of class IV are innovations on analogy to other strong classes (see below).

Class VI and VII are often characterized by an /a/ rather than /e/ root vowel, and utilize two different strategies for differentiating between the principal parts (van Coetsem 1994: ch. 8). Class VI verbs show /o:/ in the preterit singular and plural, and the same, short *a*-grade form for the participle as in the present. Class VII verbs show the same syncretism between the present and participle forms, but mark the preterit singular and plural by partial reduplication. Except for *s*+stop clusters, Gothic reduplicates only the initial consonant(s) and a fixed short /e/, as in \*flōkan ‘bewail’ ~ <faiflok> containing reduplicated [fe-], ‘they bewailed’, with fixed vowel and simplified onset cluster. With *s*+stop clusters, we find reduplication of *st-* and *sk-*, as in *gastaldan* ‘to receive’ ~ *gastaistald* (Nehemiah 5:16) and *afskaidan* ‘to part ways’ ~ *afskaiskaidun* (Luke 9:33). Class VII verbs are not attested as a productive tense-marking strategy in North or West Germanic (Bammesberger 1986: 64, Ringe 2006: 248–250).<sup>3</sup> Still, while the *e*-grade and *a*-grade strong verbs mark the principal parts differently, *a*-grade class VII verbs look strikingly like class I–III verbs of the *e*-grade in their consonants, and class VI verbs are built similarly to class IV and V (van Coetsem 1994: 127).

Returning to the origin of /ē/ in the preterit class IV and V, the structural similarity to class IV verbs suggests analogous use of a long vowel to distinguish tense. The reconstructed zero-grade from *geban* ‘give’ would be \*gbum in the preterit plural, not a licit consonant cluster in Germanic; a lengthened root vowel on analogy to the class VI verbs repairs an otherwise unpronounceable consonant sequence. Beyond repair strategies, analogy must also account for the presence of /ē/ in IV verbs like *stelan* ~ *stēlum* ‘steal ~ stole.PL.’ and *k<sup>w</sup>eman* ~ *k<sup>w</sup>ēmum* ‘come ~ came.PL.’, the expected forms of which would have been *stelan* ~ \*stulum and *k<sup>w</sup>eman* ~ \*k<sup>w</sup>umum, since zero-grade liquids and nasals develop epenthetic /u/ in Germanic, preserved in the preterit presents. This is also the case with class III verbs *helpan* ~ *hulpum* and *drinkan* ~ *drunkum*. Analogy does not occur since class III verbs are structurally similar to class VII, which does not use vowel length alternations to distinguish principal parts.

In addition to the 7 strong classes, weak verbs – see §4 – formed their preterits with a /d/ or /t/ suffix, as in Modern Eng. *-ed* or Modern Germ. *-te*. This “dental preterit” was organized into four classes, which often correspond to particular functional distinctions: (i) causatives/denominatives, (ii) iteratives, (iii) duratives and (iv) inchoatives.

The trend is a steady expansion of weak verbs at the expense of the strong (cf. Lieberman et al. 2007, Carroll et al. 2012). The process has left many residues and much variation at any given stage. Historically strong class III OEng. *meltan* ‘melt’ has become weak (e.g. *melted*) but preserves a frozen adjectival form, *molten*, derived from the strong paradigm. Ongoing change is visible in Modern Germ. *backen* ‘bake’, which has an archaic strong preterit *buk* alongside the more common, innovative *backte* in the 1/3 sg., but the strong past participle *gebacken*. Movement in the other direction is rarer,

e.g. American English *dive*, historically weak, has developed a strong preterit *dove*, likely on analogy to verbs such as *drive* ~ *drove*.

A last category is the so-called preterit presents, verbs with a preterit form but present meaning. This category includes what have become our modern modal verbs, e.g. OHG *mugan* ‘may’, *skulan* ‘shall’, *wellen* ‘want’; others like *muozan* ‘must’, *kunnan* ‘can’, *durfan* ‘be allowed to’ (ONor. *þurfa* ‘require, need’, Modern Dutch *durven* ‘dare’). Early on, these arguably acted more like full verbs. These maintain archaisms from the strong paradigms, including the syncretism of the first and third person singular. Additionally, since the preterit presents already have a strong preterit form in the present, the past is formed by adding the weak dental suffix.

### Nominal system

The IE system of case, number and gender is complex relative to most attested languages, including Germanic. IE had 8 cases: nominative, accusative, dative, genitive, ablative, locative, instrumental and vocative. Germanic is reconstructed as having six, based on attestations in the earliest texts. The early daughters retain four robustly: Nominative, Accusative, Dative and Genitive, and two more in limited fashion: Vocative is attested only in Gothic, and Instrumental is attested in West Germanic. The receding of the Vocative and Instrumental are early examples of case loss. Vocative was absorbed into other classes; even in Gothic it was already identical in form to the Nominative in the plural, a pattern inherited from PIE. Vocative differed in form from the nominative singular only when the nominative singular ended in *-s*, which includes only Gothic masculine *a*-stems (IE short *o*-stems) and their subclasses. In that event, the vocative was identical in form to the accusative singular (*wulfs* ‘wolf’ NOM.SG.; *wulf* ‘wolf’ ACC./VOC.SG. An exceptional case is the *u*-stem *sunus* ‘son’ NOM.SG.; *sunu* ‘son’ ACC./VOC.SG.

The instrumental was attested in Old High German but restricted to a subset of nominal classes even early, specifically the masculine and neuter *a*-stems, the masculine *u*-stems and the masculine and feminine *i*-stems. Some of the nouns belonging to these classes are particularly well suited to instrumental use, as in *swërt* ‘sword’ (neuter *a*-stem), *scaft* ‘spear’ (feminine *i*-stem) and *witu* ‘wood’ (masculine *u*-stem, OEng. *wudu*), as in (16) from the Hildebrandslied, l. 40. Such an example is provided below, in (17), from the Hildebrandslied l. 37, as Haðubrant expresses his distrust of his opponent (his father, Hildebrand).

- (16) *wili*                      *mih*                      *dinu*                      *speru*                      *werpan*  
 want.2SG.PRET.OPT.    me.ACC.SG.    your.INSTR.    spear.INSTR.SG.    throw  
 ‘You would want to kill (lit. ‘throw’) me with your spear’

Haðubrant’s utterance three lines earlier shows an instrumental – doubly marked morphologically and by a preposition, *mit* ‘with’ – as he suggests how the untrusted Hildebrand should hand over an armband as a gift.

- (17) *mit*                      *geru*                                      *scal*                      *man*                      *geba*                      *infahan*  
 with    spear(point).INSTR.SG.    should    one                      gifts                      receive  
 ‘One should receive gifts with the point of the spear.’

Eventually the dative incorporates the function of the instrumental, most often also marked with a preposition. This is likely accelerated by the propensity for heavy *a*-stems in particular to appear without inflectional morphology, due to both their syllable structure and analogy to uninflected nominative and accusative forms in the paradigm (Braune & Eggers 1987: 182–183).

Number is similar to IE, with singular, plural and dual. The dual was restricted in the earliest attestations. While dual pronouns exist in Old Norse, Old English, Old Saxon and Gothic, only Gothic retains a verbal paradigm specific to the dual; in other languages, dual pronouns inflect as plural (Howe 1996).

Gender largely echoes what is reconstructed for (late) IE. Germanic retains three genders – masculine, feminine and neuter – but daughters tend to lose this distinction in the plural. While many early languages maintained gender-specific inflections in the nominal paradigm (e.g. for a group of men), variation increases when the plural involves nouns of more than one gender. For example, Luke 8:1–4 (Codex Argenteus) describes a gathering of both men and women who listen to Jesus preach. Following a dative plural (which is ambiguous for all genders in the pronoun and strong adjectives), the relative pronoun for the gathering of men and women is masculine plural *þaiei: gaqumanaim þan hiumam managaïm jah þaim þaiei us baurgim gaiddjedun du imma, qab þairh gajukon:* (‘And when many people were gathered together, and were come to him out of every city, he spake by a parable.’, Luke 8:4, www.wulfila.be). Masculine is sometimes the default for indefinites (cf. Codex Argenteus, Luke 8:13), but also appears as a masculine plural relative pronoun *þaiei*, even anaphorically to a feminine singular noun *so managei* ‘the multitude’, *alja so managei, þaiei ni kunnun witob, fraqïþanai sind* ‘but this multitude who do not know the law are cursed’ (Codex Argenteus, John 7:49). Over time, the trend is to retain gender distinctions in the singular but to collapse gender into a unified plural in the nominal and pronominal paradigms (except Modern Icelandic; Einarsson 1945).

The nominal system remains based on inherited classes, organized primarily around thematic vowels and secondarily around gender. Key Germanic noun classes include the *a*-stems (masculine and neuter, IE *o*-stems), *ō*-stems (feminine, IE *ā*-stems), *i*-stems (masculine and feminine), *u*-stems (masculine and feminine) and subclasses of these. Consonant stems include the *r*-stems (a small class, familial relations), *n*-stems (the basis for new weak adjective endings) and *nd*-stems (based on participles). Additionally, there are smaller classes like the *-izl/-az* stems (originally comprising names for baby animals; this class over time comes to play a major role in German plural formation, along with *i*-stems).

Masculine *a*-stems, illustrated with ‘day’ in Table 8.5, are one of the largest and most stable noun classes. Still, changes are visible even early in the textual evidence of the collapse of some parts of the paradigm, due to a combination of sound change and (subsequent) morphological changes.

**TABLE 8.5 A-STEM MASCULINE NOUNS** (adapted from Prokosch 1939: 241–242; Braune & Eggers 1987: 182)

	Germanic	Gothic	Old Norse	Old High German	Old Saxon	Old English
<b>singular</b>						
nom.	*dagaz	<i>dags</i>	<i>dagr</i>	<i>tag</i>	<i>dag</i>	<i>dæg</i>
acc.	*daga(n)	<i>dag</i>	<i>dag</i>	<i>tag</i>	<i>dag</i>	<i>dæg</i>
gen.	*dagas(a)	<i>dagis</i>	<i>dags</i>	<i>tages (-as)</i>	<i>dages, -as</i>	<i>dæges</i>
dat.	*dagai	<i>daga</i>	<i>degi</i>	<i>tage (-a)</i>	<i>dage, -a</i>	<i>dæge</i>
instr.	*dagō	—	—	<i>tagu, -o</i>	<i>dagu, -o</i>	—
voc.	*dag	<i>dag</i>	—	—	—	—
<b>plural</b>						
nom.	*dagōzez	<i>dagōs</i>	<i>dagar</i>	<i>taga (-ā)</i>	<i>dagos, -as, -a</i>	<i>dagas</i>
acc.	*daganz	<i>dagans</i>	<i>daga</i>	<i>taga (-ā)</i>	<i>dagos, -as, -a</i>	<i>dagas</i>
gen.	*dagō(n)	<i>dagē</i>	<i>daga</i>	<i>tago</i>	<i>dago</i>	<i>daga</i>
dat.	*dagamaz	<i>dagam</i>	<i>ðogum</i>	<i>tagum, -om, -un, -on</i>	<i>dragon, -un</i>	<i>dagum</i>
instr.	*dagamiz?	—	—	—	—	—
voc.	?*dagōzez	—	—	—	—	—

Gothic maintains a comparatively unambiguous paradigm (except the vocative), while later daughters exhibit changes that introduce ambiguity or reduce uniformity across the paradigm. For example, West Germanic languages collapse the nominative and accusative singular, either through the extension of the accusative to the nominative, or due to the loss of the nominal inflection *-s* (in Gothic) or *-r* (in Old Norse). A similar collapse is visible in the nominative and accusative plural, though Old English and Old Saxon extend the nominative into the accusative, while Old High German extends the accusative into the nominative. In terms of sound change, Old Norse shows *i*-umlaut in dative singular, and *u*-umlaut in the dative plural, resulting in three different vowels across the paradigm. Old English, on the other hand, experienced Anglo-Frisian fronting of /a/ to /æ/ (“brightening”), visible in the singular but reversed in the plural, by height harmony with the suffixal /a/; the dative plural change was the same process. This split introduces an alternation correlated with the singular/plural distinction.

The increasing tendency to mark singular/plural within the existing nominal class was widespread across Germanic. Similarly, nouns realign from one noun class to another or adopt new inflections based on formal similarities to other classes. Consider the Old High German comparison below, between masculine *i*-stem *gast* ‘guest’, neuter *-iz/-az* stem *lamb* ‘lamb’ and neuter *a*-stem *wort* ‘word’.

**TABLE 8.6 EXAMPLES OF OTHER NOUN CLASSES** (adapted from Sonderegger 2003; Prokosch 1939: 245; Braune & Eggers 1987: 182–198)

	<i>gast</i> ( <i>i</i> -stem)		<i>lamb</i> ( <i>iz/az</i> -stem)	<i>wort</i> ( <i>a</i> -stem)	
	oldest OHG	late OHG	OHG	oldest OHG	late OHG
<b>singular</b>					
nom.	<i>gast</i>	<i>gast</i>	<i>lamb</i>	<i>wort</i>	<i>wort</i>
acc.	<i>gast</i>	<i>gast</i>	<i>lamb</i>	<i>wort</i>	<i>wort</i>
gen.	<i>gastes</i>	<i>gastes</i>	<i>lambes</i>	<i>wortes</i>	<i>wortes</i>
dat.	<i>gaste</i>	<i>gaste</i>	<i>lambe</i>	<i>worte</i>	<i>worte</i>
instr.	<i>gastu/gastiu</i>	—	<i>lambu, -o</i>	<i>wortu</i>	—
<b>plural</b>					
nom.	<i>gasti</i>	<i>geste</i>	<i>lambir</i>	<i>wort</i>	<i>wort</i>
acc.	<i>gasti</i>	<i>geste</i>	<i>lambir</i>	<i>wort</i>	<i>wort</i>
gen.	<i>gasteo, -io</i>	<i>gesto</i>	<i>lambiro</i>	<i>worto</i>	<i>worto</i>
dat.	<i>gastim</i>	<i>gesten</i>	<i>lambirum, -un, -on</i>	<i>wortum</i>	<i>worten</i>

Through sound changes, especially the rise of umlaut in plurals, the clearest difference in the paradigm is between singular and plural, marked by a vowel alternation originally conditioned by suffixal *-i* (*gast*, oldest OHG); this remains as a morphologically conditioned vowel alternation after the conditioning high vowel has reduced to a schwa and disappeared. The same alternation arises for *lamb*. The neuter *a*-stems, like *wort*, have more paradigmatic ambiguity, with uninflected forms for nominative and accusative, in both the singular and plural. At the same time, the inflectional endings in the singular are identical to those of the masculine *i*-stems and of the *-iz/-az* stems. In late Old High German and especially in Middle High German, some former *a*-stems mark plurality with umlaut, on analogy to unrelated noun classes, in order to repair a deficient paradigm. Crucially, this involves the spread of a portmanteau affix from one class as a plural marker in another. Some of these include MHG *rat* ~ *reder* ‘wheels’, *blad* ~ *bleter* ‘leaves, pages’

and *hūs ~ hiuser* ‘houses’. This is indicative of the tendency toward marking of number to expand at the expense of noun classes (*Pluralprofilierung*; Luiten 2011, Luiten et al. 2013, and many others), and reflects the weakening of the inherited noun class system.

One additional noun class deserves mention, *n*-stems, from which the so-called weak adjectival endings were derived. Not inherited from IE, weak adjectival endings were employed in Germanic to mark definiteness (Behaghel 1923: 197); the use of inherited, strong adjectival endings marked indefinites. In short, definiteness marking in Germanic arose with an innovative system of complementary adjectival paradigms before the development of definite or indefinite articles (Harbert 2007: 130ff.). For pronouns, the fullest available historical treatment is Howe 1996.

## WORD FORMATION

Trips (2014: 385), focusing on Germanic and noting some exceptions, writes this about diachronic derivational morphology:

Much has been written about derivation, and more generally about word formation, from a theoretical and synchronic perspective describing and analyzing instances of derivation in many different contemporary languages. Much less attention has been paid to historical aspects of derivation (or word formation), especially to the rise of affixes and systematic descriptions and analyses of affixal inventories available in diachronic stages of languages.

We briefly survey some key developments in nominal and verbal derivation diachronically.

Reflecting syntactic patterns, compounding is restricted by the relative ordering of constituents. Modern West Germanic dialects and languages exhibit both left-headed and right-headed structures, meaning that various grammatical elements either precede or follow their complements. For example, prepositions precede the nouns they modify; in English, verbs precede their complement – often a noun – but the opposite is true for German and Dutch. Reflecting the underlying typology of the respective systems as either left- or right-headed, English and German may have more or less free combinations of adjectives, nouns and verbs in the formation of nominal compounds. However, a preposition must uniformly precede its complement, even in word formation, since postpositional prepositions are a mere subset of prepositions as a whole (e.g. *underworld*, German *Unterseeboot* ‘submarine’; lit. ‘undersea boat’ (König & Gast 2012: 260ff.). Additionally, verb-object compounds – even in Modern English – are no longer productive. Compounds such as *turncoat*, *cutthroat* and *spendthrift* sound archaic. Object-verb constructions like *barnburner* and *fishfry* are acceptable for most speakers, and the strategy is productive for new compounds.

Germanic compounding is more rigid, being exclusively right-headed, reflecting the syntax of the early Germanic period (Harbert 2007: 30ff.). Runic and Gothic compounds of two nouns reflect older “stem-based” compounds, in which the first element is not a stand-alone noun, marked only with a root and thematic noun class marker; case is assigned only to the rightmost element in the compound (Lass 1994: 194). Possible remnants of these earlier stages include such examples as *Schafskopf* (a card game, literally ‘sheep’s head’), which shows an archaic word-internal genitive inflection on the first element of the compound, a paradigmatic *Fugenelement* ‘linking element’. A non-paradigmatic *-s* is obligatory in some noun-noun compounds where not historically present, such as when the first element is feminine, e.g. *Schönheitsideal* ‘ideal of beauty’ (cf. *die Schönheit*). Duden



(2006: 721–723) suggests all linking elements are at least homonyms of original (genitive) inflectional endings; linking elements generally no longer serve an inflectional purpose but rather are primarily phonologically conditioned and characterized by considerable regional variation. Evidence of earlier stages is also reflected when the first element is a weak noun (from the earlier *n*-stem class): *Wangenbein* ‘cheekbone’, cf. *die Wange*; *Zungenspitze* ‘tip of the tongue’, cf. *die Zunge*; *nach Herzenslied* ‘to your heart’s content’, cf. *das Herz*.

In the verb system, word formation through the affixation of prefixes and suffixes affected both the syntax and morphology of Germanic. Kuroda’s (2007) study of verbal prefixes in Otfrid’s *Evangelienbuch* shows the development of productive verbal prefixes by Old High German, which were built on basic verb forms: though the basic verb form may only have taken a prepositional phrase as complement, verbal prefixes *bi-*, *fir-*, *int-*, *ir-*, *missi-*, *ver-* and *zir-* supplement the valence of the basic form to allow for accusative objects; this is also possible even when the basic verb is intransitive (2007: 31). Perhaps more striking are instances where verbal prefixes modify the meaning of the verb, and “[d]er Bedeutungsunterschied, der zwischen dem Basisverb und dem Kompositum besteht, ist jedoch so drastisch, dass dazwischen ein Ableitungsverhältnis kaum angenommen werden kann” (the difference in meaning which exists between the base verb and the compound is, however, so drastic that a derivational relationship between them can hardly be assumed; 2007: 41). This trend toward greater verbal prefixation is attested in Old High German and increases in frequency toward the modern period.

A well-known Germanic derivational pattern is the already-noted rise of weak verbs, marking the past tense by a dental suffix (Hill 2004, 2010, Ringe 2006). Four classes, based at least originally on the semantic properties or grammatical origin of the weak verb, are often labeled *i-iv* (Prokosch 1939: 154–159). Class *i* is “causatives” formed by adding a *-jan* suffix to an existing lexical item. They are often derived from the preterit singular (*a*-grade) forms of strong verbs, or from existing nouns, i.e. Gothic *mats* ‘meat, food’ > *matjan* ‘to eat’. This yields a variety of minimal pairs that have persisted into the modern languages, e.g. Modern Germ. *trinken* ‘drink’ > *tränken* ‘water [a horse]’ (Modern Eng. *drink* ~ *drench*), *sitzen/setzen*, *liegen/legen*, *essen/ätzen*, etc. Class *i* verbs were also built from adjectives, as in Goth. *daups* ‘dead’ yielding *daupjan* ‘kill’; or OHG *frewide* ‘happiness’ yielding *frewen* (from \**frawjan*) ‘make happy’.

Class *ii* verbs are often denominatives, often iterative. An example is OHG *salbōn*, ‘anoint, apply salve’. This class is also exemplified by such OHG verbs as *dionōn* ‘serve’ and *samanōn* ‘gather’. Class *iii* is typically durative, e.g. OHG *habēn* ‘have’, *lebēn* ‘live’ and *wonēn* ‘reside’, all denoting events occurring over stretches of time. Still, some variability is possible between class *ii* and *iii* verbs, especially when the verb is ambiguously iterative or durative, as in OHG *hazzōn/hazzēn* ‘hate’, *klagōn/klagēn* ‘complain’, *fastōn/fastēn* ‘fast (i.e. to abstain from eating)’ and *wisōn/wisēn* ‘search for something; visit’ (Braune & Eggers 1987: 299). On the one hand, that some verbs jump between class *ii* and *iii* implies that these classes were semantically transparent to some degree still in Old High German; on the other hand, confusion regarding class assignment may also suggest that distinctions between classes were disappearing. This was surely further complicated by the reduction of the unstressed vowels that distinguished verb classes (see above).

The fourth class has a \*-*nan* suffix and inchoative meaning (Prokosch 1939: 156–157). It was a small class, not productive in West Germanic. Goth. *gadaupnan* ‘die’, *fraquistnan* ‘perish’ and Old Norse *wakna* ‘awaken’ convey a change of state, while Goth. *fullnan* ‘become full’ and *usmēnnan* ‘be proclaimed’ express the beginning of a process. Already disappearing as a morphologically marked verbal class from many of the daughter languages in the earliest texts – including notably West Germanic – the syntactic function of this class is later subsumed by the periphrastic passive.

The weak preterit's origin is debated, possibly deriving from an IE past participle suffix \*-to- or a grammaticalized form of IE \*d<sup>h</sup>eh<sub>1</sub>-/d<sup>h</sup>oh<sub>1</sub>- 'put'. One view reconstructs its origin in reduplicated plural forms attested in Gothic, which look like a grammaticalized form of an IE strong preterit; the lack of reduplicating forms in later-attested North and West Germanic – excepting OEng. *dyde* > Modern Eng. *did* – results from haplology (Ringe 2006). Arguing that the North and West Germanic non-reduplicated suffixes are more indicative of the proto-form, Hill (2004, 2010) reconstructs the origins of the dental preterit using an IE imperfect, which “secondarily adopted the optative inflections of the corresponding reduplicated perfect”; the differences between North and West Germanic, on the one hand, and Gothic, on the other, result from divergent leveling of the Germanic paradigm (2010: 448–451).

## SYNTAX

We now turn to interactions across syntax, morphology and phonology (interfaces) with attention to grammaticalization and ongoing changes: word order and verb second, complementizers, definiteness and changes in tense-mood-aspect.

### Word order

Syntactic reconstruction is one of the most contested areas of linguistic prehistory (Ferraresi & Goldbach 2008 and now Walkden 2014). Still, the predominant view is that Germanic maintained inherited IE verb-final word order, and German and Dutch maintain some verb-final constructions throughout history (e.g. subordinate clauses, impersonal imperatives). However, from the earliest attestations in Runic and Gothic, fronting/raising of finite verbs occurs, pragmatically and syntactically conditioned. Arguably under way before attested texts (Wackernagel 1892), changes in the position of the finite verb are interwoven with related developments in phonology and morphology. Going along with these broad and gradual changes are numerous decidedly Germanic innovations in finite verb position: the inherited IE clause-final position of the verb gradually yields to competing word orders, first in conditioned environments, and later obligatorily. Variable, pragmatically conditioned word order is gradually replaced by increasingly restrictive word orders. Stockwell (1977) proposes a unified model, including the earliest, verb-final word order inherited from IE, through the V2 constraint (verb second; the obligatory appearance of the finite verb as the second syntactic constituent) observable in German and Dutch, to the SVO word order of English and Scandinavian (in most contexts). Stockwell's account largely represents the mainline thinking regarding the impetus and descriptive typology for the change in position of the finite verb in Germanic.

Inherited SOV is found in Runic, though verb fronting was possible. Variation occurs even within inscriptions, even though the corpus is tiny: only 34 early inscriptions contain a verb and another element that would allow us to determine word order patterns (Antonsen 1975: 24). The R6 Runestone shows both the inherited clause-final position of the verb, along with the innovative earlier verb placement (Antonsen 2002).

- (18) *ek hrزاز/hrapaz satido [s]tain[a] . . . swabaharjaz sairawidaz*  
 I Hrazaz/Hrapaz set stone.ACC.DEF. . . Swabaharjaz with wide wounds . . .

. . . *stainawarijaz fahido*  
 . . . Stainawarijaz painted

‘I Hrazaz/Hrapaz set the stone . . . Swabaharjaz with-wide-wounds . . . Satainawarijaz painted.’<sup>24</sup>

Clauses in which the finite verb appears earlier than the expected clause-final position arguably show a degree of emphasis and suggest a degree of flexibility in verb position based on pragmatics. The verb can also appear earlier in the clause based on syntactic context, e.g. in the presence of negation or a *wh*-element, or in commands (Thiersch 1978, Eythórsson 1995, 1996). Eythórsson (1995: 26, 249–257) notes that fronting of the finite verb of this specific type (“V-to-C” movement – movement of the verb from a lower syntactic position to a higher verbal/clausal position in the clause) occurs in these contexts in all old Germanic languages. “CP operants” – e.g. *wh*-elements and negation, which activate the highest syntactic phrase not otherwise present in the earliest-attested Germanic languages – provide a syntactic environment that triggers a non-clause-final word order, showing that innovative positions of the finite verb in Germanic may be conditioned grammatically, as well as pragmatically. For all old Germanic languages except Gothic and Old English, raising of the finite verb is also obligatory when nominal complements are topicalized (Eythórsson 1995: 3, 22–28), providing supporting evidence that obligatory fronting later supplanted the optional fronting of the finite verb. Despite positive evidence of variable word order in Runic, however, the small corpus prevents broader generalizations about the frequency of orders in early Germanic. Still, variation suggests that change was under way from at least the earliest written evidence (200–500 CE). Given that IE is reconstructed as having the finite verb in a clause-final position, it is likely that the older SOV order appears alongside a newer V2 (verb second) or SVO (subject, verb, object) order attested in subsequent varieties.

### Rise of V2

The V2 constraint is present in main clauses in German and Dutch, and attested also for other historical periods of other Germanic languages, including Old English and Old Norse (Faarlund 2004). That V2 developed from an earlier SOV (verb-final) order in Germanic is generally accepted, though the process is much more nuanced. In fact, this change remains incomplete in some modern daughter languages, with German and Dutch employing SOV in subordinate clauses.

### V2 over time

Despite evidence of verb fronting, the position of the finite verb is not fixed in early Germanic. West Norse (Old Icelandic, Old Norwegian) and East Norse (Old Swedish, Old Danish, Old Gutnish) exhibit V2 in both main and subordinate clauses, though “[d]eviations from this [in the *Edda*] are shown to be due to the poetic character of the text (although it is possible that certain patterns reflect a more archaic stage at which verb-second was not obligatory)” (Eythórsson (1995: 190–191). In West Germanic, V2 is by no means obligatory in all contexts, with Old High German texts also exhibiting verb-initial and V3 word orders. In the *Isidor*, Axel identifies a number of instances where the verb raises to a medial position, even above a complementizer in subordinate clauses (2007: 79–81). Robinson (1994: 356) notes, also in the *Isidor*, that the position of the finite verb is also often not consistent with the Latin *Vorlage*. Multiple verb-raising strategies were indigenous to Old High German, with V2 the most common: “the translator of the Old High German *Isidor* imposed an apparently native verb-second word order . . . despite the example of the Latin original, and also of apparent exceptions to this principle where the verb appears later in the clause” (Robinson 1994: 356). In declarative and main clauses, the Old High German translation agrees with the Latin *Vorlage* on 135

occasions, though the Old High German position of the verb differs from the Latin text in 255 clauses, including 205 V2 clauses imposed by the Old High German translator (Lippert 1974). One account for this is variation in the landing site of the finite verb, that is, where the verb appears in the clause, if not in clause-final position. Axel (2007) argues that V3 may result from the raising of the verb to a medial position higher than the clause-final position of IE but lower than the ultimate V2 position attested in Modern German (see Weerman 1989 for a similar argument for Old English). Commonly argued for Old English is that a poverty of OV structures relative to VO structures in the input resulted in an unraveling of the inherited, underlying OV word order (Stockwell 1977 and others). A corresponding lack of SVOV structures in Old English concurrently destabilized the transition from OV to V2 predicted by Stockwell's (1977) typology, resulting ultimately in an (S)VVO order (Fuß & Trips 2002). Differences in the position of the finite verb in Old Norse, Old High German and Old English provide evidence that the position of the finite verb was highly variable, both across the earliest attested Germanic languages and within the grammars of the individual varieties. The subsequent adoption of V2 structures was a parallel development, occurring in West Germanic centuries after the split from North Germanic.<sup>5</sup>

Another account for word order variation relates to the structure of the IE clause and subsequent changes in Germanic. As Kiparsky (1995: 140) notes, Germanic maintained the structure of the IE complementizer phrase (CP), which had both a topic and a focus position above C, the highest syntactic position to which a finite verb could raise. Even assuming a fixed position for the finite verb, at C, variation in the position of any number of constituents would then result in V1, V2 or V3 word order, depending on whether there were zero, one or two elements preceding the finite verb. Remnants of this earlier clausal structure are illustrated by Eythórsson (1995: 101–102), who provides one example from the Old Norse *Völundarkviða* and three from Gothic where *wh-* elements appear to the left of C, as well as additional rare instances in Old Norse and Gothic of an indefinite negative quantifier appearing to the left of a negated verb (1995: 145–149). In view of the rarity of these examples, the poetic nature of the Old Norse texts, and the possible influence of the Greek *Vorlage* on Gothic, Eythórsson notes – and we concur – that these exceptional clausal structures are likely archaisms reflecting earlier stages that allowed for greater flexibility of word order. The later collapse of these two clausal positions through their generalization as a single topic position reduced the number of possible word orders, eliminating the possibility of V3 constructions and increasing the propensity toward V1 or V2 structures characteristic of German and Dutch, respectively in imperative and conditional clauses and in V2 main clauses. This collapse of topic and focus positions was a prerequisite for obligatory V2, primarily in Continental West Germanic.

### Clausal syntax

As noted, Runic and Gothic exhibit syntactically conditioned verb raising in the presence of a CP-operant – a *wh*-element, negative particle – or in imperative constructions. Verb raising is possible but not obligatory elsewhere pragmatically – an asymmetry suggesting ongoing change toward the development of subordinating constructions in Germanic not present in (P)IE. As noted in Matasović (p. 166), (P)IE did not have subordinating conjunctions but rather relied on subordinating strategies using participles.<sup>6</sup> However, all branches of Germanic use neuter demonstratives as complementizers (Goth. *þat*, OHG *dass/dazs*, OEng. *þæt*, ONor. *að*, *at*, *att*) (Harbert 2007: 417). Likely derived from IE uninflected deictic particles, Germanic complementizers independently develop toward

innovative subordinating structures. As in the example below (Otfrid's *Evangelienbuch*), the neuter demonstrative *thasz* may have been parsed as the direct object of the first clause but later reanalyzed as the head of the subordinate clause.

- (19) Otfrids *Evangelienbuch*: Ev. 2, 2, 8

[<sub>CP</sub> *Joh gizalta in sār thasz*] [<sub>CP</sub> *thiu sālida untar in uuas*].

[<sub>CP</sub> *Joh gizalta in sār*] [<sub>CP</sub> *thasz thiu sālida untar in uuas*].

‘And he told them immediately that good fortune was among them’

The potential ambiguity is yet greater, given that both V2 and verb-final constructions are possible in main clauses in (early) Old High German; the maintenance of verb-final constructions in subordinate clauses in Modern German and Modern Dutch is an archaic remnant of an earlier stage in the language.

Related to this class of neuter demonstratives as complementizers are relative pronouns, which similarly build subordinating structures not present in IE. The differences between bare relatives and relative particles in Goth. *ei*, Old Norse *es*, *er* and OEng. *þe*, *þē*, *ðe* attest to an innovative rather than inherited characteristic. Moreover, these differences across the branches hint at a parallel post-*Ausgliederung* development or, more likely, at earlier dialectal variation within common Germanic. Subsequent developments in all the Germanic languages transitioned from the use of relativizing particles to use of demonstrative pronouns as relative pronouns (Lehmann 2005–2007). Demonstrative and relative pronouns in Germanic derive predominantly from IE deictic elements, *t*, *h* and *s*, a marked difference from Romance and other languages that derive relative pronouns from *wh*-elements (IE \**k<sup>w</sup>o*-).<sup>7</sup> Gothic shows a possible intermediate stage, where relativizing pronouns co-occur as clitics on demonstrative pronouns. In the example below *ei* serves as a stand-alone relativizing particle; as a relativizing enclitic particle attached to a demonstrative pronoun, as in *þatei*; and also as a relativizing enclitic particle attached to a demonstrative pronoun marked for case, number and gender, i.e. serving as a relative pronoun.

- (20) Codex Argenteus, John 12:9

*fanh þan manageins filu Iudaie þatei Iesus jainar ist, jah qemun, ni in Iesus ainis,*  
found then multitude many Jews.GEN.PL. REL Jesus there is and came.3PL. not for  
Jesus only

*ak ei jah Lazaru sehaina, þanei urraisida us dauþaim.*

but **rel** also Lazarus see.PRET.OPT.3PL. **rel.M.ACC.SG.** raised from death

‘Then a great multitude of Jews found that Jesus was there, and they came, not only for Jesus, but also that they might see Lazarus, whom he raised from the dead.’  
(Our translation)

Marking for case, number and gender on subordinating elements was apparently typologically rare in IE daughters and crosslinguistically (p. 166–167, this volume), but here it appears to have arisen independently many times across the family.

## Definites

Reconstructions of Germanic do not posit definite articles, but definiteness – with case, number and gender – was encoded with suffixes based on noun class. However, by the

earliest attestations, demonstratives derived from deictic elements *þ*, *h* and *s* appear. Though *þ* occurs across the family, variation in the pronominal paradigm once again points to either dialectal differences in Germanic or parallel developments across branches. One oft-mentioned similarity between Old Norse and Gothic is that, despite a majority of pronouns and demonstratives with *þ*, archaic Gothic *sa* (NOM.SG.M.DEM.) and *sō* (NOM.SG.F.DEM.) correspond to Old Norse *sá* (NOM.SG.M.DEM.) and *sú* (NOM.SG.F.DEM.). Second, pronouns and demonstratives with *h* are attested mostly in North Sea Germanic. The use of demonstratives, however, is similar across the family, and early texts employ demonstratives not only in emphatic or contrastive contexts but also anaphorically, as in the OHG gospel harmony *Tatian*:

(21) *Tatian*, chapter 1, sentence 1

*In anaginne uuas uuort inti thaz uuort*

in beginning.DEF was word.DEF.SG.M.A-STEM and that (same) word.DEF.SG.M.A-STEM  
*uuas mit gote inti got selbo uuas thaz uuort.*

was with God and God self was that (same) word.DEF.SG.M.A-STEM.

‘In the beginning was the word, and that same word was with God, and God (him) self was that word.’

The first use of *uuort* ‘word’ employs the older definiteness marking, with an unprecedented noun characteristic of the Latin *Vorlage*. However, in the second and third uses, it is preceded by the demonstrative *thaz*, where the Latin *Vorlage* lacks a demonstrative: *In principio erat verbum et verbum erat apud deum et deus erat verbum*. Old High German seems to allow unprecedented bare nouns, on the one hand, but also employs preceded nouns in anaphoric contexts; a conservative translation of *thaz* in this case would be ‘that same’. This provides evidence of a gradual transition from the inherited IE system, moving toward the innovative precursors of definite articles.

In terms of diachrony, definite articles are not reconstructed for Germanic, or even attested in the early daughter languages. Rather, demonstrative pronouns are reanalyzed as definite articles, through language-internal processes of grammaticalization. Formalizing the process, van Gelderen (2007, 2009, 2011) categorizes grammaticalization of demonstrative pronouns as the loss of the feature [+emphasis], with the corresponding retention of interpretable phi-features (or [i-phi]) marking for person and number agreement.<sup>8</sup> Arguing that “the real sources of change are internal principles that bias the learner toward certain structures” (van Gelderen 2011: 4), van Gelderen motivates this reanalysis through the Head Preference Principle (HPP) (2007: 283; 2011: 13–14). The HPP holds that specifiers (and phrases) such as demonstratives are less preferable than syntactic heads such as definite articles, and that the natural process of language change involves reanalyzing the former as the latter. Such a change in the grammar is consistent with Kuryłowicz’s (1965: 52) definition of grammaticalization as “the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status”.

Indefinite articles are similarly unattested in the earliest Germanic languages. Indefinite nouns are marked by strong adjectival inflection preceding the noun. Indefinite articles developed from the numeral ‘one’, with a corresponding reanalysis from number to indefinite marker. Much like the development of definite articles, this was a process of grammaticalization, advancing to a more grammatical function. Because the numbers one through four inflected in Gothic,<sup>9</sup> the numeral ‘one’ already inflected for case, number and gender. Grammaticalization, then, involved reanalysis of a quantifier as an article (van Gelderen 2011: 202).<sup>10</sup>

### Tense, aspect, mood

The Germanic systems of tense, aspect and mood (TAM) reflect a rather drastic departure from IE, with subsequent development of indigenous innovations. Relative to IE, Germanic underwent leveling followed by the rise of new tenses and aspects. The development of aspectual auxiliaries as part of a cyclical pattern – even as a small portion of ongoing change – is observable over a period of centuries in a variety of languages, including the earliest attestations of Old English and Polish through to the modern languages (van Gelderen 2011: 273–276).

Reconstructions of IE often posit a four-way system of aspect and tense, differentiating perfect and imperfect, and past and present, respectively; the present tense likely may have expressed future tense<sup>11</sup> (p. 162, this volume). Germanic collapses these two systems, using the perfect of strong verbs to mark past, while the present tense forms continued to mark both present and future tense. Perfective aspect may be expressed through use of verbal prefixes, most commonly *ga-* (derived from IE \**kom*, ‘together, with, collective, intensive’). Inherently perfective verbs often carry this prefix, as in *gadauþnan* ‘die’ or *gatairan* ‘destroy’ (Modern Germ. *zerreißen* ‘to rip, to tear (apart)’), both only attested with the perfective prefix *ga-*. Still, that the perfective prefix was productive is apparent in common verbs such as Goth. *waurljan* ‘to prepare, bring about’, *saihwana* ‘to see’, *hausjan* ‘to hear’ and *gaggan* ‘to go’, which have perfective variants *gawaurkjan*, *gasaihwana*, *gahausan* and *gagaggan*. This perfective prefix may also serve in some cases as a kind of aspectual perfect marker, as in the opening line of the 9th century Hildebrandslied (22).

- (22) *Ik gihorta ðat seggen*  
 I heard.1SG.PERF. that said  
 ‘I have heard it told . . .’

The narrator addresses the listener/reader directly, using a first person pronoun. The perfective *gihorta* establishes the narrator’s authority, having heard the tale prior to the present narrative; indeed, the story may have been well known, with analogues across Germanic languages, as well as in Persian, Russian and Old Irish (Baesecke 1940). Without a dedicated pluperfect, however, ordering of events in the past tense is often accomplished using participles.

- (23) Codex Argenteus, Mark 5:39  
*jah innatgaggands qap du im: hwa auhjob jah gretip þata barn*  
 and enter.PRES.PTCP. said.3SG. PRO to them why fuss.2PL. and cry.2PL. PRO that child  
*ni gadauþnoda ak slepip*  
 not died.3.SG.PRET.PERF but sleeps  
 ‘And when he had come in, he said to them: why do you fuss and weep? The girl is not dead, but sleeps.’ (Our translation)

While a literal translation of *jah innatgaggands qap* might be ‘and entering (he) said’ or ‘and as he entered, he said’, we draw from context – and from frequent attestation of present participles elsewhere in the Gothic Bible – that these events are sequential rather than simultaneous; coordinated structures employ *jah* ‘and’.

A similar example shows the use of a present participle to order events at the beginning of the verse, with the act of exiting (*usgaggands*) preceding subsequent events.

Additionally, the perfective prefix *ga-* is used to convey something that had already happened prior to the moment of speaking, in *gasahu*.

(24) Codex Argenteus, Luke 1:22

*usgaggands þan ni mahta du im rodjan jah froþun þammei siun*

EXIT.PRES.PTCP. then not might PRO to them speak and perceived.2PL. that.REL vision  
*gasahu in alh;*

3SG.SEEN.PERF. in temple

‘When he came out, he could not speak to them, and they perceived that he had seen a vision in the temple.’ (Our translation)

The second half of the verse exemplifies an emerging periphrastic construction, combining a past form of *wisan* ‘be’ and participial *bandwjands* ‘signaling’.

(25) Codex Argenteus, Luke 1:22

*jah silba was bandwjands im jah <ga>was dumbs.*

and same was signaling them.DAT.PL. and was.1SG. speechless.

‘... and the same one [Zacharias] was signaling to them, and [he] was speechless.’

The English translation may be most colloquial with a present progressive, but the Gothic participle is not unambiguously verbal: the inflectional *-s* could also be a nominative suffix (masculine *nd*-stem), that is, a copular construction between the grammatical subject and an inflected nominal or adjectival form meaning, here, ‘one who signals; a signaling one’. This interpretation of otherwise ambiguous Gothic data is consistent with the later development of the periphrastic perfect, in which verbs meaning ‘have’ and ‘be’ become auxiliaries and co-occur with participles. Crucially, these participles frequently appear with inflectional suffixes, supporting the argument that they were originally nominal (or adjectival) rather than verbal. Perfective prefixes *ga-*, *gi-* and *ge-* (IE \*kom-) regularly co-occur with nominal suffixes but are notably absent from reflexes of double modal constructions and more grammatical uses of OHG *lāzzan* ‘let, allow’, *heizzan* ‘command’ and *werdan* ‘become’. The reanalysis of these as past participles corresponds with the loss of inflectional affixes.

Germanic also reduced other parts of the TAM system. IE is reconstructed with – in addition to the indicative – an optative and subjunctive that collapse into a single “optative” in Germanic (Matasović 2012). Marked in Old High German by verbal affixes – *ē* in the present and *ī* in the preterit – the optative begins to collapse with unstressed syllable reduction. In the present indicative, the distinction between present indicative and present optative is lost, first in the first person plural, e.g. *faremēs* ‘we go’. Later, it is lost throughout the paradigm as full vowels reduced and centralized, and unstressed long syllables shortened, i.e. OHG *faru* ‘I go.1SG.PRES.IND.’ and OHG *fare* ‘I go.1SG.PRES.OPT.’ both collapse to MHG *fare*. The distinction is maintained only in verb classes where other sound changes caused vowel mutations, e.g. MHG *nime* ‘take.1SG.PRES.IND.’ but *neme* ‘take.1SG.PRES.OPT.’ The preterit optative is identical to the preterit indicative for weak verbs in Old High German, and the OHG *-i* suffix for strong verbs causes umlaut before reducing; umlaut demarcating the optative (German Subjunctive I) is thereafter morphologically conditioned. Ultimately, as in other parts of the grammar, morphological marking of aspect yielded to periphrastic constructions.

The other component of the innovative Germanic periphrastic perfect is the development of auxiliaries from the verbs ‘have’ and ‘be’. OHG *habēn* (or *eigan*) was originally



a lexical verb, e.g. ‘to own’. To ‘have’ something that is in a current (perfective) state implies that it was previously not in that state; thus, the current state is interpreted as a present tense, albeit one derived from an earlier completed action. For example, the oft-cited OHG sentence *phigboum habeta sum giflanzotan* (Tatian) is arguably to be interpreted as ‘someone had planted a fig tree’ (Lockwood 1968: 115). Some skepticism would be warranted, though, since the inflected participle looks more like an adjectival form than a perfective verbal form. Still, while somewhat ambiguous in its contemporary reading, such an argument is a plausible example of change under way.

## CONCLUSION

Across the phenomena that separate Germanic from IE and separate the daughters from one another, we see surprising cohesion and perseverance in the kinds of change and patterns found today:

- vocalic and obstruent chain shifts,
- case loss,
- strong verbs becoming weak,
- independent development of typologically rare subordinating strategies.

From IE to the first textual evidence, Germanic languages include phonetic/phonological reductions and various drastic reorganizations:

- shift away from nominal classes toward a coherent singular/plural distinction,
- expansion of periphrastic verbal forms to create marking for tense, aspect and mood,
- development of non-inherited inflectional affixes and inflectional categories that increase the complexity of the inflectional system.

Questions remain about the sociolinguistic contexts in which these changes took place. Extended contact would be consistent with the sorts of morphological reduction found: increasingly fixed word order and collapse of inherited TAM marking. Some innovations may reflect prehistoric variation:

- the development of three innovative strategies for marking preterit (classes VI and VII, weak classes),
- differences in the weak preterit plural ending, separating Gothic from other languages,
- wide variation in the pronominal inventories,
- a tripartite division in subordinating conjunctions and relative particles (Gothic *ei*, ONor. *es*, *er* and OEng. *þe*, *þē*, *ðe*),
- divergent chronologies of nasal loss and compensatory lengthening, cf. *mouth* ~ *Mund*.

Still, the Germanic family portrait shows profound familial similarities, not only in inherited forms and structures, but also in parallel developments among sometimes distant daughters; and shared developments and borrowings during extended, repeated periods of contact.

Progress today is rapid thanks to developments in general, Germanic and IE linguistics, including better understanding of crosslinguistic patterns of language change (e.g. how common glide hardening turns out to be), new models of how change works (e.g. in prosody) and newer issues (the rise of complementizers).

## FURTHER READING

Bammesberger, Alfred 1986, *Der Aufbau des germanischen Verbalsystems*, Winter. Heidelberg; 1990, *Die Morphologie des urgermanischen Nomens*, Winter, Heidelberg

Detailed treatments of derivation and inflection of early Germanic nouns and verbs.

Jacobs, Neil 2005, *Yiddish: A Linguistic Introduction*, Cambridge University Press, Cambridge

Yiddish is underdiscussed in the context of early Germanic but provides unique perspectives in terms of language contact and dialectal relationships.

Lehmann, Winfred P. 1967, *A Reader in Nineteenth-Century Historical Indo-European Linguistics*, Indiana University Press, Bloomington

Edited collection of seminal texts in IE linguistics, with commentary on historical context within the field.

Nielsen, Hans Frede 2000, *The Early Runic Language of Scandinavia*, Winter, Heidelberg

Delivers the best survey of the *Ausgliederung* and internal structure of the Germanic family, contextualized in external evidence.

Prokosch, Eduard 1939, *A Comparative Germanic Grammar* (William Dwight Whitney Linguistic Series), Linguistic Society of America–University of Pennsylvania, Philadelphia

Classic older work on Germanic in the IE context.

Ringe, Don 2006, *From Proto-Indo-European to Proto-Germanic*, Oxford University Press, Oxford

The clearest statement of Germanic sound and word form patterns and developments from an IEist perspective, with a few pages on derivational morphology and mention of syntax and lexicon.

Robinson, Orrin W. 1992, *Old English and Its Closest Relatives: A Survey of the Earliest Germanic Languages*, Stanford University Press, Stanford

The most reliable and accessible introduction to the family.

## NOTES

- 1 Many glottalic proposals posit implosives rather than voiced stops, and implosion is a form of “hypervoicing”, phonetically enhancing a [voice] contrast.
- 2 Dresher and Lahiri (1991: 263) suggest that this restriction may in fact have been limited to word-initial feet. If so, similar patterns might be shared across dialects.
- 3 Possible remnants of reduplication in West Germanic are controversial, such as OEng. *heht* from Gmc \**haihait*, or Modern Eng. *do* ~ *did*, possibly related to \**tōn* (Braune & Eggers 1987: §381–382).
- 4 While *fahido* is related to the term for ‘to paint’, it is also used broadly as ‘to fashion runes’, which may also include carving.

- 5 The development of V2 is parallel but not independent: verb fronting strategies show that change was under way when the branches diverged.
- 6 Gothic makes extensive use of participles, though this relates more directly to the tense/aspect system than to subordinating structures (§5.6).
- 7 Some modern languages use *wh*-elements as relative pronouns in restricted contexts, most notably Eng. *who/whom* and *which*, in competition with null complementizers and *that*. Germ. *was* and Dutch *wat* serve as relative pronouns for indefinite or inanimate nouns; Germ. *wo* occurs dialectally as a relative pronoun for animate nouns, and Dutch *wie* or *welke* may also be used, in formal or archaic contexts.
- 8 Since many demonstrative pronouns already mark number and gender, this is a clear case of grammaticalization, cf. Gothic demonstratives *sah* (M. SG.), *þatuh* (N. SG.), *sōh* (F. SG.) and *þāih* (M. PL.; other genders not attested); OEng. *se* (M. SG.), *þæt* (N. SG.), *sēo* (F. SG.), *þā* (PL. all genders), etc.
- 9 Some other IE daughters – including Slavic languages – also decline numbers.
- 10 The nature of this reanalysis “depends on one’s assumptions about what numerals are and where indefinite articles are in the DP” (van Gelderen 2011: 202); we are concerned here only with the broader picture.
- 11 Alternatively, the IE subjunctive may fulfill this function.

## REFERENCES

- Antonsen, Elmer 1975, *A Concise Grammar of the Older Runic Inscriptions*, Niemeyer, Tübingen
- Antonsen, Elmer 2002, *Runes and Germanic Linguistics*, Mouton de Gruyter, Berlin
- Árnason, Kristján 2011, *The Phonology of Icelandic and Faroese*, Oxford University Press, Oxford
- Avery, Peter & Idsardi, William 2001, “Laryngeal dimensions, completion and enhancement,” in: Hall, T.A. (ed.), *Distinctive Feature Theory*, Mouton de Gruyter, Berlin: 41–70
- Axel, Katrin 2007, *Studies on Old High German Syntax: Left Sentence Periphery, Verb Placement and Verb-second*, Benjamins, Amsterdam
- Axel, Katrin & Weiß, Helmut 2011, “Pro drop in the history of German: From Old High German to the modern dialects,” in: Wrátil, Melani & Peter Gallmann (eds.), *Empty Pronouns*, Mouton De Gruyter, Berlin: 21–51
- Baesecke, Georg 1940, “Die Indogermanische Verwandtschaft des Hildebrandliedes,” *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen. Phil. – hist. Klasse* 3: 139–153
- Bammesberger, Alfred 1986, *Der Aufbau des germanischen Verbalsystems*, Winter, Heidelberg
- Bammesberger, Alfred 1990, *Die Morphologie des urgermanischen Nomens*, Winter, Heidelberg
- Bayer, Josef 1984, “COMP in Bavarian syntax,” *The Linguistic Review* 3: 209–274
- Behaghel, Otto 1923, *Deutsche Syntax: Eine geschichtliche Darstellung*, Winter, Heidelberg
- Biberauer, Theresa & Roberts, Ian 2008, “Cascading parameter changes: Internally-driven change in Early and Middle English,” in: Eythórsson, Thórhallur (ed.), *Grammatical Change and Linguistic Theory: The Rosendal Papers*, Benjamins, Amsterdam: 79–114
- Bousquette, Joshua 2014, “Complementizer Agreement in eastern Wisconsin: (Central) Franconian features in an American heritage language community,” *Sprachtypologie und Universalienforschung / Language Typology and Universals* 67: 561–588

- Braune, Wilhelm & Eggers, Hans 1987 (2004), *Althochdeutsche Grammatik*, Niemeyer, Tübingen
- Buccini, Anthony F. 1995, "Onstaan en vroegste ontwikkeling van het Nederlandse taal-landschap," *Taal en Tongval* 48: 8–66
- Carroll, Ryan, Svare, Ragnar & Salmons, Joseph 2012, "Quantifying the evolutionary dynamics of German verbs," *Journal of Historical Linguistics* 2: 153–172
- Catford, J.C. 2001, "On Rs, rhotacism and paleophony," *Journal of the International Phonetic Association* 31: 171–185
- De Vogelaer, Gunther 2010, "Morphological change in continental West Germanic: Towards and analogical map," *Diachronica* 27: 1–31
- Dresher, B. Elan & Lahiri, Aditi 1991, "The Germanic foot: Metrical coherence in old English," *Linguistic Inquiry* 22: 251–286
- Duden, Konrad 2006, *Duden Band 4: die Grammatik*, Wermke, Matthias, Kathrin Kunkel-Razum & Werner Scholze-Stubenrecht (eds.), Bibliographisches Institut & Brockhaus, Mannheim
- Einarsson, Stefán 1945, *Icelandic*, The Johns Hopkins Press, Baltimore
- Emonds, Joseph Embley & Faarlund, Jan Terje 2014, *English: The Language of the Vikings*, Palacký University, Olomouc [Available at <http://anglistika.upol.cz/fileadmin/kaa/emonds/vikings2014.pdf>]
- Enger, Hans-Olav 2013, "Inflectional change 'sound laws' and the autonomy of morphology: The case of Scandinavian case and gender reduction," *Diachronica* 30: 1–26
- Eythórsson, Thórhallur 1995, *Verbal Syntax in the Early Germanic Languages*, Cornell University [PhD dissertation]
- Eythórsson, Thórhallur 1996, "Functional categories, cliticization and word order in the Early Germanic Languages," in: Thráinsson, Höskuldur; Epstein, Samuel D. & Peter, Steve (eds.), *Studies in Comparative Germanic Syntax*, vol. II, Kluwer, Dordrecht: 109–139
- Faarlund, Jan Terje 2004, *The Syntax of Old Norse*, Oxford University Press, Oxford
- Faarlund, Jan Terje 2012, "Linguist makes sensational claim: English is a Scandinavian language," *ScienceDaily* 2012, November 27, [www.sciencedaily.com/releases/2012/11/121127094111.htm](http://www.sciencedaily.com/releases/2012/11/121127094111.htm)
- Ferraresi, Gisella 2005, *Word Order and Phrase Structure in Gothic*, Peeters, Leuven
- Ferraresi, Gisella & Goldbach, Maria (eds.) 2008, *Principles of Syntactic Reconstruction*, Benjamins, Amsterdam
- Fertig, David 2000, *Morphological Change up Close: Two and a Half Centuries of Verbal Inflection in Nuremberg*, Niemeyer, Tübingen
- Forster, Peter, Tobias Polzin & Arne Röhl. 2006, "Evolution of English basic vocabulary within the network of Germanic languages," in: Forster, Peter & Renfrew, Colin (eds.), *Phylogenetic Methods and the Prehistory of Languages*, McDonald, Institute Cambridge: 131–138
- Fuß, Eric 2004, "Diachronic clues to pro-drop and complementizer agreement in Bavarian," in: Fuß, Eric & Trips, Carola (eds.), *Diachronic Clues to Syntactic Grammar*, Benjamins, Amsterdam: 51–100
- Fuß, Eric 2005, *The Rise of Agreement: A Formal Approach to the Syntax and Grammaticalization of Verbal Inflection*, Benjamins, Amsterdam
- Fuß, Eric & Carola Trips 2002, "Variation and change in Old and Middle English: On the validity of the Double Base Hypothesis," *Journal of Comparative Germanic Linguistics* 4: 171–224

- Goeman, Ton 1997, "De zeldzaamheid van comp-agreement in taaltypologisch en historisch opzicht. Voorkomen buiten de Germania en datering voegwoordcongruentie-vormen in het Nederlands," in: Hoekstra, Eric & Smits, C. (eds.), *Vervoegde voegwoorden: Lezingen gehouden tijdens het Dialectsymposion 1994*, Meertens, Amsterdam: 87–111
- Green, D.H. 1998, *Language and History in the Early Germanic World*, Cambridge University Press, Cambridge
- Hall, T.A. 2014, "The analysis of Westphalian German Spirantization," *Diachronica* 31: 223–266
- Halle, Morris 1997, "On stress and accent in Indo-European," *Language* 73: 275–313
- Hamp, Eric P. 1990, "The Pre-Indo-European Language of Northern (Central) Europe," in: Markey, T.L. & Greppin, John A.C. (eds.), *When Worlds Collide: Indo-Europeans and Pre-Indo-Europeans: The Bellagio Papers*, Karoma, Ann Arbor: 291–310
- Hanna, Patrizia Noel Aziz 2009, "Jespersen's Cycle and the issue of prosodic 'weakness'," in: Alexiadou, Artemis; Hankamer, Jorge; McFadden, Thomas; Nuger, Justin & Schäfer, Florian (eds.), *Advances in Comparative Germanic Syntax*, Benjamins, Amsterdam: 197–217
- Harbert, Wayne 2007, *The Germanic Languages*, Cambridge University Press, Cambridge
- Haugen, Einar 1976, *The Scandinavian Languages: An Introduction to their History*, Faber and Faber, London
- Heather, Peter 2010, *Empires and Barbarians: The Fall of Rome and the Birth of Europe*, Oxford University Press, Oxford
- Henton, Caroline; Ladefoged, Peter & Maddieson, Ian 1992, "Stops in the world's languages," *Phonetica* 49: 65–101
- Hill, Eugen 2004, "Das germanische Verb für ‚tun‘ und die Ausgänge des germanischen schwachen Präteritums," *Sprachwissenschaft* 29: 257–303
- Hill, Eugen 2010, "A case study in grammaticalized inflectional morphology: Tracing back the Germanic weak preterit," *Diachronica* 27: 411–458
- Honeybone, Patrick 2002, *Germanic Obstruent Lenition: Some Mutual Implications of Theoretical and Historical Phonology*, University of Newcastle upon Tyne [PhD dissertation]
- Howe, Steven 1996, *The Personal Pronouns in the Germanic Languages*, de Gruyter, Berlin
- Howell, Robert B. 1991, *Old English Breaking and its Germanic Analogues*, Niemeyer, Tübingen
- Howell, Robert B. & Salmons, Joseph C. 1997, "Umlautless residues in Germanic," *American Journal of Germanic Linguistics & Literatures* 9: 83–111
- Iverson, Gregory K. & Salmons, Joseph C. 1995, "Aspiration and laryngeal representation in Germanic," *Phonology* 12: 369–396
- Iverson, Gregory K. & Salmons, Joseph C. 1996, "The primacy of primary umlaut," *Beiträge zur Geschichte der deutschen Sprache und Literatur* 118: 69–86
- Iverson, Gregory K. & Salmons, Joseph C. 2003, "Laryngeal enhancement in early Germanic," *Phonology* 20: 43–72
- Iverson, Gregory K. & Salmons, Joseph C. 2011, "Final devoicing and final laryngeal neutralization," in: van Oostendorp, Marc, Ewen, Colin, Hume, Beth & Rice, Keren (eds.), *Companion to Phonology*, Volume III, Wiley-Blackwell, Oxford: 1622–1643
- Iverson, Gregory K. & Salmons, Joseph C. 2012, "Parasitic rule loss in Norse umlaut," *Journal of Germanic Linguistics* 24: 101–131
- Kathol, A. 2001, "Syntactic categories and positional shape alternations," *Journal of Comparative Germanic Linguistics* 3: 59–96

- Kerswill, Paul 2002, "Koineization and accommodation," in: Chambers, J.K., Trudgill, Peter & Schilling-Estes, Natalie (eds.), *Handbook of Language Variation and Change*, Blackwell, Oxford: 669–702
- Kiparsky, Paul 1995, "Indo-European origins of Germanic syntax," in: Battye, Adrian & Roberts, Ian (eds.), *Clause Structure and Language Change*, Oxford University Press, Oxford: 140–169
- Kiparsky, Paul & Halle, Morris 1977, "Towards a reconstruction of the Indo-European accent," in: Hyman, Larry (ed.), *Studies in Stress and Accent*, Department of Linguistics, University of Southern California: 209–238
- Kluge, Friedrich 2011, *Etymologisches Wörterbuch der deutschen Sprache*, 25th edn, de Gruyter, Berlin
- Koivulehto, Jorma 2001, "Etymologie und Lehnwortforschung: Ein Überblick um 2000," *Finnisch-ugrische Forschungen* 56: 42–78
- König, Ekkehard & Gast, Volker 2012, *Understanding English-German Contrasts*, 3rd edn, Erich Schmidt Verlag, Berlin
- Krause, Wolfgang 1971, *Die Sprache der urnordischen Runeninschriften*, Winter, Heidelberg
- Kristoffersen, Gjert 2011, "Quantity in Old Norse and peninsular North Germanic," *Journal of Comparative Germanic Linguistics* 14: 47–80
- Kroonen, Guus 2013, *Etymological Dictionary of Proto-Germanic*, Brill, Leiden
- Kuroda, Susumu 2007, "Zur valenzmodifizierenden Funktion der Verbalpräfixe im Althochdeutschen und Gegenwartsdeutschen," *Sprachwissenschaft* 32: 29–75
- Kuryłowicz, Jerzy 1965, "Zur Vorgeschichte des germanischen Verbalsystems," in: Isačenko, A. V.; Wissmann, W. & Strobach, H. (eds.), *Beiträge zur Sprachwissenschaft, Volkskunde und Literaturforschung: Wolfgang Steinitz zum 60. Geburtstag*, Akademie, Berlin: 242–247
- Küspert, Klaus-Christian 1988, *Vokalsysteme im Westnordischen: Isländisch, Färöisch, Westnordwegisch: Prinzipien der Differenzierung*, Niemeyer, Tübingen
- Kylstra, A.D. 1991, *Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen*, Rodopi, Amsterdam
- Labov, William 1994, *Principles of Linguistic Change*. Volume 1, *Internal Factors*, Blackwell, Oxford
- Lass, Roger 1994, *Old English*, Cambridge University Press, Cambridge
- Lehmann, Winfred P. 1986, *A Gothic Etymological Dictionary*, Brill, Leiden
- Lehmann, Winfred P. 2005–2007, *A Grammar of Proto-Germanic*, Jonathan Slocum (ed.), Linguistics Research Center, University of Texas, Austin
- Lieberman, Erez, Michel, Jean-Baptiste, Jackson, Joe, Tang, Tina & Nowak, Martin A. 2007, "Quantifying the evolutionary dynamics of language," *Nature* 449: 713–716
- Lippert, Jörg 1974, *Beiträge zu Technik und Syntax althochdeutscher Übersetzungen*, Fink, Munich
- Lockwood, W.B. 1968, *Historical German Syntax*, Clarendon, Oxford
- Luiten, Tyler 2011, *Old High German Nominal Morphology Revisited: Fresh Look at Old Paradigms*, University of Wisconsin – Madison [PhD dissertation]
- Luiten, Tyler, Menz, Andrea, Bagwell, Angela, Frey, Benjamin, Lindner, John, Olson, Mike, Speth, Kristin & Salmons, Joseph 2013, "Beyond the handbooks: A quantitative approach to analysis of Old High German phonology and morphology," *Beiträge zur Geschichte der deutschen Sprache und Literatur* 135: 1–18
- Mailhammer, Robert & Vennemann, Theo (eds.) 2015, *Die sprachlichen Wurzeln Europas – Linguistic Roots of Europe: Ursprung und Entwicklung – Origin and Development*, Tusculum, Copenhagen

- Matasović, Ranko 2012, "Areal typology of PIE: The case for Caucasian connections," *Transactions of the Philological Society* 110: 283–310
- McWhorter, John H. 2002, "What happened to English?," *Diachronica* 19/2: 217–272
- Meillet, Antoine 1967, *The Indo-European Dialects*, Translated by Samuel Rosenberg. University of Alabama Press, Tuscaloosa
- Menz, Andrea 2010, *Accent Type and Language Change in Germanic and Baltic-Finnic*, University of Wisconsin – Madison [PhD dissertation]
- Mortensen, David 2012, "The emergence of obstruents after high vowels," *Diachronica* 29: 434–470
- Müller, Stefan 2007, *Zum Germanischen aus laryngaltheoretischer Sicht*, de Gruyter, Berlin
- Murray, Robert, & Vennemann, Theo 1983, "Sound change and syllable structure in Germanic phonology," *Language* 59: 514–528
- Nielsen, Hans Frede 2000, *The Early Runic Language of Scandinavia*, Winter, Heidelberg
- Oftedal, Magne 1952, "On the origin of the Scandinavian tone distinction," *Norsk Tidsskrift for Sprogvidenskap* 16: 201–225
- O'Neil, Wayne 1978, "The evolution of the Germanic inflectional systems: A study in the causes of language change," *Orbis* 27: 248–286
- Orel, Vladimir 2003, *A Handbook of Germanic Etymology*, Brill, Leiden
- Page, Raymond I. 2001, "Frisian runic inscriptions," in: Munske, Horst Haider (ed.), *Handbuch des Friesischen* (with Nils Århammar; Volkert F. Faltings; Jarich Hoekstra; Oebele Vries; Alastair G.H. Walker & Ommo Wilts), Niemeyer, Tübingen: 523–530
- Polomé, Edgar C. 1972, "Germanic and the other Indo-European languages," Kufner, Herbert & van Coetsem, Frans (ed.), *Toward a Grammar of Proto-Germanic*, Niemeyer, Tübingen: 43–69
- Polomé, Edgar C. 1989, "Substrate lexicon in Germanic," *NOWELE* 14: 53–73
- Prokosch, Eduard 1939, *A Comparative Germanic Grammar* (William Dwight Whitney Linguistic Series). Linguistic Society of America–University of Pennsylvania, Philadelphia
- Reis, Marga 1985, "Satzleidende Strukturen im Deutschen: Über comp, Haupt, und Nebensätze, w-Bewegung und die Doppelkopfanalyse," in: Abraham, Werner (ed.), *Erklärende Syntax des Deutschen*, Narr, Tübingen: 271–311
- Riad, Tomas 2014, *The Phonology of Swedish*, Oxford University Press, Oxford
- Ringe, Don 2006, *From Proto-Indo-European to Proto-Germanic*, Oxford University Press, Oxford
- Roberge, Paul T. 2010, "Contact and the history of the Germanic languages," in: Hickey, Raymond (ed.), *Handbook of Language Contact*, Blackwell, Oxford: 406–431
- Robinson, Orrin W. 1992, *Old English and its Closest Relatives: A Survey of the Earliest Germanic Languages*, Stanford University Press, Stanford
- Robinson, Orrin W. 1994, "Verb-first position in the Old High German 'Isidor' translation," *Journal of English and Germanic Philology* 93: 356–373
- Salmons, Joe 1992, *Accental Change and Language Contact*, Stanford University Press, Stanford
- Salmons, Joseph C. 2004, "How (non-)Indo-European is the Germanic Lexicon? . . . And what does that mean?," in: Hyvärinen, Irma, Kallio, Petri & Korhonen, Jarmo (eds.), *Etymologie, Entlehnungen und Entwicklungen: Festschrift für Jorma Koivulehto zum 70. Geburtstag*, Société Néophilologique, Helsinki: 311–321

- Salmons, Joseph 2012, *A History of German: What the Past Reveals about Today's Language*, Oxford University Press, Oxford
- Salmons, Joseph 2015, "Language shift and the Indo-Europeanization of Europe," in: *Mailhammer & Vennemann* 2015: 147–169
- Salmons, Joseph & Smith, Laura Catharine 2005, "On the status of the Indo-European labiovelar stops," *Indogermanische Forschungen* 110: 86–96
- Schirmunski, Viktor M. 2010, *Deutsche Mundartforschung*, Akademie, Berlin [Second edition Larissa Naiditsch, Lang, first published 1962]
- Schulte, Michael 2006, "Oral traces in Runic epigraphy: Evidence from older and younger inscriptions," *Journal of Germanic Linguistics* 18: 117–151
- Sievers, Eduard 1876, *Grundzüge der Lautphysiologie: Zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*, Breitkopf & Härtel, Leipzig
- Sihler, Andrew 1995, *New Comparative Greek and Latin Grammar*, Oxford University Press, Oxford
- Smith, Laura Catharine 2004, *Cross-Level Interactions in West Germanic Phonology and Morphology*, University of Wisconsin – Madison [PhD dissertation]
- Smith, Laura Catharine & Ussishkin, Adam 2014, "The role of prosodic templates in diachrony," in: Honeybone, Patrick & Salmons, Joseph (ed.), *The Oxford Handbook of Historical Phonology*, Oxford University Press, Oxford: 262–285
- Somers, Katerina 2011, "The introduction and extension of the *-st* ending in Old High German," *Journal of Germanic Linguistics* 23: 141–181
- Somers Wicka, Katerina 2009, *From Phonology to Syntax: Pronominal Cliticization in Otfrid's Evangelienbuch*, Niemeyer, Tübingen
- Sondererger, Stefan 2003, *Althochdeutsche Sprache und Literatur: Eine Einführung in das älteste Deutsch; Darstellung und Grammatik*, Walter de Gruyter, Berlin
- Stearns, Jr., McDonald 1978, *Crimean Gothic: Analysis and Etymology of the Corpus*, Anima Libri, Saratoga
- Stockwell, Robert P. 1977, "Motivations for exbraciation in Old English," in: Li, Charles N. (ed.), *Mechanisms of Syntactic Change*, University of Texas Press, Austin
- Thiersch, Craig L. 1978, *Topics in German Syntax*, MIT, Cambridge [PhD dissertation]
- Trips, Carola 2014, "Derivation and historical change," in: Lieber, Rochelle & Stekauer, Pavol (eds.), *Oxford Handbook of Derivational Morphology*: 384–406
- Trudgill, Peter 2010, *Investigations in Sociohistorical Linguistics*, Cambridge University Press, Cambridge
- van Coetsem, Frans 1994, *The Vocalism of the Germanic Parent language*, Winter, Heidelberg
- van Gelderen, Elly 2007, "Definiteness cycle in Germanic," *Journal of Germanic Linguistics* 19: 275–308
- van Gelderen, Elly (ed.) 2009, *Cyclical Change*, Benjamins, Amsterdam
- van Gelderen, Elly 2011, *The Linguistic Cycle: Language Change and the Language Faculty*, Oxford University Press, Oxford
- Vennemann, Theo 1988, *Preference Laws for Syllable Structure and the Explanation of Sound Change*, Mouton de Gruyter, Berlin
- Verner, Karl 1875, "Eine Ausnahme der ersten Lautverschiebung," *Zeitschrift für vergleichende Sprachforschung* 23: 97–130
- Wackernagel, Jakob 1892, "Über ein Gesetz der indogermanischen Wortstellung," *Indogermanische Forschungen* 1: 333–436



- Walkden, George 2014, *Syntactic reconstruction and Proto-Germanic*. New York. Oxford University Press
- Weerman, Fred 1989, *The V2 Conspiracy: A Synchronic and a Diachronic Analysis of Verbal Positions in Germanic Languages* (Publications in Language Sciences 31), Foris, Dordrecht
- Weise, Oskar 1907, "Die sogenannte Flexion der Konjunctionen," in: Heilig, Otto & Lenz, Philipp (eds.), *Zeitschrift fuer Deutsche Mundarten* vol. 2, Franz Steiner Verlag, Stuttgart: 199–205
- Weiß, Helmut 2005, "Inflected complementizers in continental West Germanic dialects," *Zeitschrift für Dialektologie und Linguistik* 72: 148–166
- Weiß, Helmut (forthcoming), "The Wackernagel position and complementizer agreement – the emergence of a syntactic particularity at the left edge of the German middle field," in: Gisella Ferraresi, Gisella; Jäger, Agnes & Weiß, Helmut, (eds.), *Clause structure and word order in the history of German* (Oxford Studies in Historical and Diachronic Linguistics), Oxford University Press, Oxford
- Wiesinger, Peter 1970, *Phonetisch-phonologische Untersuchungen zur Vokalentwicklung in den deutschen Dialekten*, de Gruyter, Berlin

# ARMENIAN

*Birgit Anette Olsen*

## INTRODUCTION

In modern times Armenian is found in two varieties, Eastern Armenian, the official language of the Armenian Republic, also spoken in Nagorno Karabagh and in north-western Iran, and Western Armenian, originally mostly found in what is now the north-eastern part of Turkey, but now mainly scattered around the Armenian diaspora with considerable bilingual populations in Lebanon, Syria, Israel, France, Canada and the USA.

The first records of Armenian go back to the Old Persian Behistūn (Bisutūn) inscription by Darius the Great (about 520 BC), where we hear of the province *Armina-* and the people *Arminiya-*. Somewhat later, Ἀρμυνίη and Ἀρμένιοι are mentioned by Herodotus, while the indigenous literary tradition only goes back to the 5th century AD, the “golden age”, with the standardized Classical Armenian, also known as *grabar*, the written language. This will be the subject of the following pages. The creation of a written language was a side effect of the introduction of Christianity, which became a state religion as early as 301. According to tradition, the monk and missionary Mesrop Maštoc<sup>c</sup> invented the Armenian alphabet, consisting of 36 letters (cf. Table 9.1), with the purpose of translating the Holy Scripture into the vernacular, and around 410 the translation of the Bible was completed, the first attempts based on a Syriac original, while the final version was founded on the Greek. During the following centuries Greek influence continues with translations of theological and philosophical literature. Important original works from the golden age include Eznik’s *Elc alandoc<sup>c</sup>*, “Refutation of Sects”, Koriwn’s “Life of Maštoc<sup>c</sup>” and the historical accounts by Agathangelos and P<sup>a</sup>wstos Buzandac<sup>i</sup>. The later historian Movsēs Xorenac<sup>i</sup>, probably from around the 9th century, is important for his descriptions of pre-Christian traditions as famously illustrated by the poem of the birth of Vahagn (Iranian *Vərəθraγna-*, Vedic *Vṛtrahān-*). Apart from early stone inscriptions, the actual attestation is of a much later date than the 5th century, the earliest manuscript being the Moscow Gospel from 887. As a literary and learned language, the grabar survived well into the 19th century, and for religious purposes it is still in use.

Middle Armenian, which already shows dialectal differentiation, is mainly transmitted in the western dialect of the Cilician kingdom (1080–1375; cf. the meticulous treatment by Karst (1901)). From around the 17th century, it gradually developed into the numerous modern dialects and the two standardized literary languages of Eastern and Western Armenian.

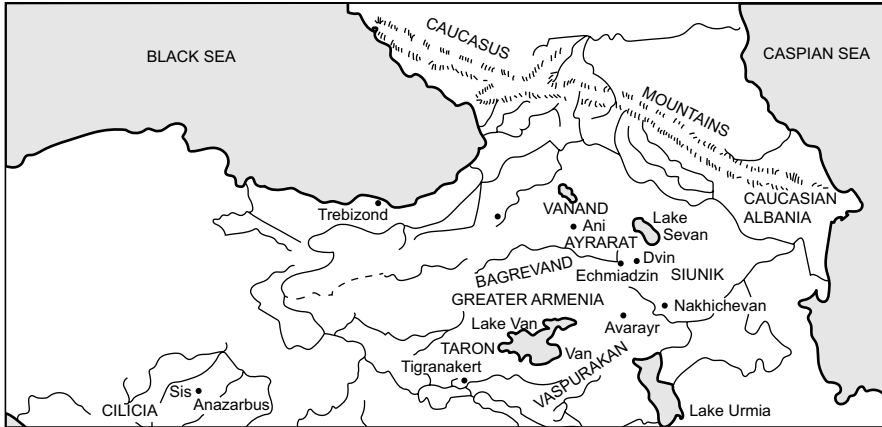
During the first millennium BC, the Armenians settled in the region around Lake Van in eastern Turkey, which was then dominated by the non-Indo-European Urartians. This is approximately the region known by the Hittites as *ḫayaša-*, which has often been compared with Arm. *hay* ‘an Armenian’, *Hayastan* ‘Armenia’. As a separate Indo-European language, Armenian is not as obviously connected to any of the other branches as Indian is to Iranian, or Baltic to Slavic. According to Herodotus, the Armenians are identified with

TABLE 9.1 THE ARMENIAN ALPHABET

	Transliteration	Value	Name		Transliteration	Value	Name
Ա	u	1	ayb	Ճ	č	100	čē
Բ	b	2	ben	Մ	m	200	men
Գ	g	3	gim	Ե	y	300	yi
Դ	d	4	da	Ն	n	400	nu
Ե	e	5	eč <sup>c</sup>	Շ	š	500	ša
Զ	z	6	za	Ո	o	600	o
Է	ē	7	ē	Չ	č <sup>c</sup>	700	č <sup>c</sup> a
Ը	ə	8	et <sup>c</sup>	Պ	p	800	pē
Թ	t <sup>c</sup>	9	t <sup>c</sup> o	Ջ	j	900	jē
Ժ	ž	10	žē	Ի	í	1000	ía
Ի	i	20	ini	Ս	s	2000	sē
Լ	l	30	liwn	Վ	v	3000	vev
Խ	x	40	xē	Տ	t	4000	tiwn
Ծ	c	50	ca	Ր	r	5000	rē
Կ	k	60	ken	Յ	c <sup>c</sup>	6000	c <sup>c</sup> o
Հ	h	70	ho	Վ	w	7000	hiwn
Ձ	j	80	ja	Փ	p <sup>c</sup>	8000	p <sup>c</sup> iwr
Ղ	ł	90	lat	Ք	k <sup>c</sup>	9000	k <sup>c</sup> ē

the Phrygians, and while this is hardly an exact description, the language does share a number of linguistic features with Phrygian, Albanian and especially Greek, which has induced some scholars, such as Klingenschmitt (1994: 244f.), Matzinger (1998: 118 and later) and Holst (2009), to connect these languages under the common heading of “Balkanindogermanisch”. Clackson, on the other hand, in a monograph dedicated to this particular question (1994), fails to recognize a significant relation even between Greek and Armenian.

The recognition of Armenian as an independent branch of the Indo-European family came only in the late 19th century. Until then it had been considered an Iranian language, and, indeed, the Iranian elements of vocabulary, word formation and composition, not to mention loan translations and syntactic features, are so abundant that this early misconception is quite understandable. The largest part consists of Parthian loans, introduced during the Arsacid dynasty (247 BC–224 AD) and exhibiting particular north-western dialectal features such as \*δ > Arm. *r*, e.g. *aroyr* ‘brass’ < \*rauða vs. South-West Iran. *y*, cf. MParth. *hwd* vs. ManMPers. *rwy* ‘copper, brass’; preservation of *-rd-*, e.g. *vard* ‘rose’ as opposed to MPers. *gul*; and \*θr > (*r*)*h*, e.g. *pah* ‘watch, ward’ < \*pāθra- vs. *s* in ManMPers. *p*’s. Occasionally a single word or stem is found in several dialectal variants. A striking example is IE \*wergóm > Iran. \*warj, borrowed from Middle Parthian as *warž* ‘study’; from Middle Persian as *vard-* in *vardapet* ‘master’; from a third Iranian dialect, possibly by Parthian mediation, as *varj* ‘reward, hire’; and finally (with analogical *o*-grade in the root) preserved in the indigenous noun *gorc* ‘work’. A systematic survey of phonetic distinctions between the Middle Parthian and the later Middle Persian loans is found in Bolognesi’s classical study (1960), followed by the useful updating by Schmitt (1983).



**MAP 9.1** ARMENIAN IN ANCIENT AND MEDIEVAL TIMES

Source: Adapted from: Walker, C. J. 1984, *Armenia. The Survival of a Nation*, Croom Helm, London–Sydney: 22

Other important sources of borrowings into Armenian are Greek and Syriac, both intimately connected with the introduction of Christianity, but apart from such easily recognizable elements, the Armenian vocabulary abounds in more or less obscure lexical items. As observed by Solta (1990: 13), Ačařyan’s monumental etymological dictionary (1926–35) contains 10,722 head words, of which 4016 are recognized loanwords, while 5572 are registered as being of unknown origin, and only 713 are considered inherited. These bare numbers would give some idea of the challenges connected with Armenian etymology.

### Note on the transcription

Aspiration of a consonant is indicated by a raised <sup>ˆ</sup>, e.g.  $t^c = [t^h]$ ;  $c$ ,  $j$  and  $c^c$  stand for the affricates  $[t^s]$ ,  $[d^z]$  and  $[t^{sh}]$ ;  $\check{c}$ ,  $\check{j}$ , and  $\check{c}^c$  for the “shibilants”  $[t^s]$ ,  $[d^z]$  and  $[t^{sh}]$ ;  $y$  equals  $[i]$ ,  $\acute{r}$  is a geminate  $r$ , and  $l$  perhaps a retroflex lateral (cf. Schmitt 1981: 29).  $\bar{E}$  derives from a former  $i$ -diphthong,  $*ey$  or  $*oy$ , and must originally have been a long vowel; however, in the classical language, where quantitative distinctions had been eliminated, the pronunciation was probably a closed  $e$  as opposed to the more open  $e$  from the old short vowel  $*e$ . Finally, the digraph  $ow$  is pronounced and transcribed as  $u$ .

## PHONOLOGY

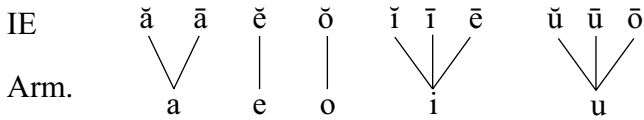
The phonology of Armenian is characterized by a series of fundamental changes, and since the inherited vocabulary is relatively limited the etymological interpretation of a given word or morpheme often depends on the precise formulation of a specific conditioned sound law. Thus, many details are still awaiting their final clarification.

### Vowels

A characteristic feature of Armenian vocalism is the loss of the quantitative oppositions before the earliest attestation of the language. While long  $\bar{a}$ ,  $\bar{i}$  and  $\bar{u}$  simply merged with

their short counterparts, and short *ě* and *ǫ* maintained an open quality, the long vowels *ē* and *ō* underwent narrowing, finally merging with *ī/ī* and *ū/ū* respectively.

We thus arrive at the following unconditioned developments in post-laryngealistic terms:



### Examples:

- \*ǎ *acem* ‘I carry, bring, conduct’, Lat. *agō*, Gr. ἄγω
- \*ā *mayr* ‘mother’, Lat. *māter*, Dor. μᾶτηρ
- \*ě *berem* ‘I carry’, Lat. *ferō*, Gr. φέρω
- \*ē *mit* ‘mind’, Gr. μῆδος
- \*ǫ *ost* ‘branch’, Gr. ὄχος, Goth. *asts*
- \*ō *tur* ‘gift’, Gr. δῶρον
- \*ī *egit* ‘(s)he found’, Ved. *ávidat*
- \*ī *c<sup>h</sup>in* ‘vulture’, Gr. ἰκτινος
- \*ū *nu* ‘daughter-in-law’, Lat. *nurus*, Gr. νύος
- \*ū *mukn* ‘mouse’, Lat. *mūs*, Gr. μῦς

Further raising of *e* to *i* and *o* to *u* occurred before the nasals *n* and *m*, cf. e.g. *hin* ‘old’ < \*senos (Lith. *sėnas*, Lat. *senex*, *senior*) and *cunr* ‘knee’ < \*ġonu (Gr. γόνυ).

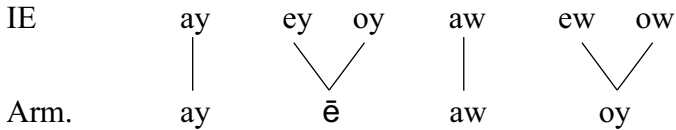
Other vocalic changes include umlaut and epenthesis. As already observed by Meillet (1903: 501), we sometimes find *-e-* for expected *-i-* when the following syllable includes *u*, e.g. *lezu* ‘tongue’ beside Ved. *jihvā-*, but apparently the rule, or at least the tendency, may be extended to cover all sequences of high vowels in consecutive syllables exhibiting a sort of “dissimilatory umlaut”, beside *i-u* > *e-u* also *i-i* > *e-i*, *u-i* > *o-i* and *u-u* > *o-u*, e.g. *korust* ‘destruction’ beside Lith. *gūrti* ‘fall to pieces’ (Olsen 1999: 801–804). The rule also applies to secondary *i* and *u* from older \**ē* and \**ō*, e.g. *tēr* ‘master’ vs. *teru<sup>h</sup>iwn* ‘domination’ rather than *tiru<sup>h</sup>iwn* (\**ē* > *i* in unstressed position).

In certain cases the vowels *i* and *u* as well as the semivowel *y* are observed to cause epenthesis, e.g. *ayl* ‘other, different’, Lat. *alius*, or *artawsr* ‘tear’ < \*dra<sup>h</sup>ku, Gr. δάκρυ. The conditioning of *i/y*-epenthesis clearly differs from that of *u*-epenthesis, and while the exact formulation of the rules may still be debated, the preceding vowel always seems to be *-a-* or *-o-*. Apparently, other decisive factors have been syllable length and stress, which are further connected with the loss of final syllables, e.g. *artawsr* vs. pl. *artasuk<sup>h</sup>*.

The status of other suggested conditioned changes affecting the vowels, such as *e* > *a* in *tasn* ‘ten’ vs. Lat. *decem*, or *o* > *a* in *ateam* ‘hate’ vs. Lat. *odium*, remains more uncertain.

### Diphthongs

After the separation of Anatolian, the protolanguage or “core Indo-European” possessed six short diphthongs, \**ay*, \**ey* and \**oy*, \**aw*, \**ew* and \**ow*. As is also the case in, e.g., Italic, the diphthongs whose first part was an *a*-vowel remained intact in Armenian as *ay* and *aw*, while the remaining *i*-diphthongs merged in *ē*, and the *u*-diphthongs in *oy*, presumably via an intermediate stage \**ow*.

**Examples:**

- \*ay *ayc* ‘goat’, Gr. αῖξ
- \*ey aor. *e-dēz* ‘(s)he piled up’ < \*e-d<sup>h</sup>eyǵ<sup>h</sup>-et, Ved. pres. *dēhmi* ‘smear’
- \*oy *dēz* ‘heap’, Gr. τοῖχος ‘wall’
- \*aw *awt<sup>c</sup>* ‘passing the night’, Gr. αῦλις ‘lodging for the night’
- \*ew *joyl* ‘molten mass’ < \*ǵ<sup>h</sup>ewtlo-, Ved. *hōtrā-* ‘libation, sacrifice’
- \*ow *p<sup>c</sup>oyt<sup>c</sup>* ‘care; haste’, Gr. σπουδή ‘haste, zeal, effort’

A similar development is found in older Iranian loanwords where *i*-diphthongs are reflected as *ē*, *u*-diphthongs as *oy*, e.g. *k<sup>c</sup>ēn* ‘hatred, animosity’, Av. *kāēnā-* ‘punishment, revenge’; *goyn* ‘colour’, Av. *gaona-*.

*Apocope, syncope and vowel weakening*

The early fixation of the stress on the penultimate in Proto-Armenian had severe consequences for the further phonetic development of vowels and diphthongs: the vowels of final syllables underwent apocope so that the synchronic accent regularly came to fall on the final syllable, e.g. \*eb<sup>h</sup>eret ‘(s)he carried’ > *ebér*. Moreover, the unstressed high vowels *i* and *u*, whatever their origin, were weakened to a mostly unwritten prop vowel, *ə*, a phenomenon generally described as “syncope”, e.g. nom. \*senos > *hin* ‘old’ vs. gen. \*senosyo > *hnoy*, nom./acc. \*dōrom > *tur* ‘gift’ vs. gen. \*dōrosyo > *troy*, or \*penk<sup>w</sup>e > *hing* ‘five’ with apocope beside *hnge-tasan* ‘fifteen’ with syncope.

Similarly, the old diphthongs (\*ey/oy >) *ē* and (\*ew/ow >) *oy* in unstressed position were continued as *i* and *u*, e.g. *sēr* ‘love’, gen. *sroy*, *p<sup>c</sup>oyt<sup>c</sup>* ‘care; haste’, gen. *p<sup>c</sup>ut<sup>c</sup>oy*, and the sequence *ea* (< *e/i* + *a* after the loss of an intervocalic consonant) as *e*, e.g. ptcp. *bereal* ‘having carried’, gen. *bereloy*.

*Vowel contractions*

After the elimination of intervocalic \*y, \*s and partly \*w and the lenition of intervocalic stops, the subsequent hiatus sequences were generally followed by vocalic contractions (cf. Schmitt 1981: 73f.). By contraction of homorganic vowels the result is the corresponding short vowel, while *e* is deleted in contact with a following or preceding *a* or *o*:

e – e	> e	*treyes > <i>erek<sup>c</sup></i> ‘three’
o – o	> o	*b <sup>h</sup> oso-g <sup>w</sup> o- > <i>bok</i> ‘barefoot’
a – a	> a	*pə <sub>2</sub> tṛb <sup>h</sup> i > <i>harb</i> ‘father’ (instr. sg.)
e – a	> a	*wesṛ- (*gear- >) <i>gar-un</i> ‘spring’
a – e	> a	*pə <sub>2</sub> teres > <i>hark<sup>c</sup></i> ‘fathers’ (nom. pl.)
e – o	> o	*swesores > <i>k<sup>c</sup>ork<sup>c</sup></i> ‘sisters’ (nom. pl.)

*Syllabic sonants*

In general, the syllabic sonants developed a preceding supporting vowel *a*, i.e. \*ɾ, \*l̥, \*ŋ, \*m̥ > *ar* (*/ar̥/*), *al* (*/al̥/*), *an*, *am*:

- \*ɾ *bard* ‘pile’, Ved. *bhṛtī-* ‘carrying, support’  
 \*l̥ *malj* ‘gall’, Av. *mərəzāna-* ‘belly’  
 \*ŋ *an-* ‘un-’, Ved. *a-*, Gr. *ἀ-*, Lat. *in-*, Goth. *un-*  
 \*m̥ *tamb* ‘ham’, Lith. *timpa* ‘sinew’, ONor. *pomb* ‘paunch; bowstring’

The same development is found in prevocalic position, i.e. mainly before an original prevocalic laryngeal, as, e.g., *am* ‘year’ < \*sm̥hah₂-, cf. Ved. *sāmā-* ‘season’.

**Consonants**

While the most striking feature of the consonant system seen in a historical perspective is certainly the Armenian sound shift, other recurrent features such as palatalizations, lenitions, metatheses and the development of consonant clusters have resulted in profound transformations of the inherited pattern.

*Stops*

By the Armenian sound shift, tenues are, in traditional terms, continued as voiceless aspirates, mediae as tenues, and voiced aspirates as mediae. However, the development of \*p usually goes a step further to *h* or zero (cf. also Klingenschmitt 1982: 165–168 for a discussion of initial *p<sup>c</sup>*- in cases like *p<sup>c</sup>lanim* ‘fall’, Lith. *pūlti*). This view has been challenged by adherents of the so-called glottalic theory, according to whom the Indo-European “mediae” were in reality glottalized stops, faithfully continued as such into modern Armenian dialects, e.g. Gamkrelidze 1985. As Armenian belongs to the “satəm” group of the Indo-European languages, the velars and labiovelars merge, at least before non-palatal vowels, while the palatals are subject to further palatalization, \*k > *s*, \*g > *c* and \*g<sup>h</sup> > *j*. Thus we arrive at the following table of unconditioned developments:

IE		Armenian	Examples
*p	>	h/Ø	*pod-ŋ > <i>otn</i> ‘foot’, Gr. acc. <i>πόδα</i> ; *pedom > <i>het</i> ‘footstep’
*t	>	t <sup>c</sup>	<i>tur</i> ‘sword’, Gr. <i>τόπος</i> ‘piercing’
*k <sup>(w)</sup>	>	k <sup>c</sup>	<i>k<sup>c</sup>erem</i> ‘scratch’, Gr. <i>κείρω</i> ‘cut’; <i>k<sup>c</sup>an</i> ‘than’, Lat. <i>quam</i>
*k̥	>	s	<i>sirt</i> ‘heart’, Lith. <i>širdis</i>
*b	>	p	<i>stipem</i> ‘press’, Gr. <i>στειβω</i> ‘step on’
*d	>	t	<i>tur</i> ‘gift’, Gr. <i>δώρον</i>
*g <sup>(w)</sup>	>	k	<i>k<sup>w</sup>unk</i> ‘crane’, Lat. <i>grūs</i> , Gr. <i>γέρας</i> ; *g <sup>w</sup> ow- > <i>kov</i> ‘cow’, Gr. <i>βοῦς</i>
*g̊	>	c	*g <sup>h</sup> onu > <i>cunr</i> ‘knee’, Skr. <i>jānu</i> , Gr. <i>γόυυ</i>
*b <sup>h</sup>	>	b	<i>berem</i> ‘carry’, Skr. <i>bhāratī</i> , Gr. <i>φέρω</i>
*d <sup>h</sup>	>	d	<i>durk<sup>c</sup></i> ‘door’, Gr. <i>θύρα</i> , Lat. <i>forēs</i>
*g <sup>(w)h</sup>	>	g	<i>geljk<sup>c</sup></i> ‘glands’, Lith. <i>gėležiūnes</i> ; *g <sup>wh</sup> - in <i>goneay</i> ‘above all’, Gr. <i>εὐθενής</i> ‘abundantly’, Lith. <i>ganà</i> ‘enough’
*g <sup>h</sup>	>	j	<i>jmeim</i> ‘winter’, OCS <i>zima</i> , Skr. <i>heman</i> ‘in winter’

After *u*, including old *u*-diphthongs, the contrast between (labio)velars and palatals is neutralized in favor of the palatal, thus \*lewkos > *loys* ‘light’, Ved. *-rocas-*, Av. *raocah-*; \*d<sup>h</sup>ugh<sub>2</sub>tēr > *dustr* ‘daughter’, Av. *dugədar-*, Lith. *duktė*.

A fourth, sporadically attested series of voiceless aspirates \*p<sup>h</sup>, \*t<sup>h</sup> and \*k<sup>h</sup>, whatever its origin, seems to have yielded Arm. *p<sup>c</sup>, t<sup>c</sup>, x*, e.g. *tup<sup>c</sup>* ‘bush’, Gr. τύφη ‘a plant used for stuffing’, *ort<sup>c</sup>* ‘calf’, Skr. *pṛthuka-*, *c<sup>c</sup>ax* ‘twig’, Skr. *śākhā* ‘branch’, *xaxank<sup>c</sup>* ‘laughter’, Gr. καγάζω ‘laugh’.

In intervocalic position, at least some of the original tenuis and voiced aspirates are subject to lenition. Here the development of the labials \*p, \*b<sup>h</sup> > -w- is secured by clear evidence, cf. \*(h<sub>1</sub>)epi > *ew* ‘and’, Skr. *āpi* ‘also’, \*-V-b<sup>h</sup>oros > -(a)wor, lit. ‘-bearer, -carrier’ (adjective and agent noun suffix), and the same goes for the palatal \*ǵ<sup>h</sup> > z, cf. \*d<sup>h</sup>eyǵ<sup>h</sup>os (or \*d<sup>h</sup>oyǵ<sup>h</sup>os) > *dēz* ‘heap’, Gr. τεῖχος ‘wall’, Av. *pairi-daēza-* ‘enclosure’. The regular continuation of intervocalic \*-t- is \*-w- before back vowels, as in \*ǵenəto > *cnaw* ‘(s)he was born’, Gr. ἐγένετο, vs. \*-y- before front vowels, as in \*pə<sub>2</sub>tēr > *hayr* ‘father’, Gr. πατήρ. However, this only seems to be the case in the position immediately after the synchronically stressed syllable. Thus, the \*-t- (dissimilated from \*-tw-) of \*k<sup>w</sup>etores *čork<sup>c</sup>* ‘four’, Ved. *catvārah*, is completely lost after the palatalization of \*k<sup>w</sup> before -e-.

The expected reflex of \*-d<sup>h</sup>- remains a matter of controversy: according to some, the outcome is -z- (e.g. \*h<sub>1</sub>lewd<sup>h</sup>- > *eluzanem*, aor. *eluzi* ‘extract, produce’, Ved. *ródhati* ‘grows’, Gr. fut. ἐλεύσομαι ‘will come’), while others assume a development to -r-, similar to that of \*-δ- in loanwords of Parthian origin (Jasanoff (1979: 145) points to the verb *ayrem* ‘burn’: Gr. αἶθω, which may, however, be an Iranian loanword). Apparently the velars and labiovelars were resistant to lenition, cf. aor. \*e-lik<sup>w</sup>et > *elik<sup>c</sup>* ‘(s)he left’, Gr. ἔλιπε (cf. an alternative scenario in Kortlandt 1980a).

As in other satəm languages, original velars and labiovelars are prone to secondary palatalization before front vowels and \*y in Armenian, though the exact conditioning is somewhat doubtful. According to an old idea by Holger Pedersen, later elaborated by Pisani (1950), Armenian and Albanian agree on a diverse treatment of simple velars and labiovelars: the labiovelar tenuis and voiced aspirate \*k<sup>w</sup> and \*g<sup>wh</sup> and the velar media \*g would undergo palatalization to č, ĵ and č respectively, while the labiovelar media \*g<sup>w</sup> (somewhat like the situation in Greek) and the plain velar tenuis and voiced aspirate \*k and \*g<sup>h</sup> would remain unaffected:

### Velars and labiovelars before front vowels and y:

*k > k <sup>c</sup>	*k <sup>w</sup> > č <sup>c</sup>
*g > č	*g <sup>w</sup> > k
*g <sup>h</sup> > g	*g <sup>wh</sup> > ĵ

Cf. \*k<sup>w</sup> > č<sup>c</sup> in \*k<sup>w</sup>etores > *čork<sup>c</sup>* vs. \*k > k<sup>c</sup> in *k<sup>c</sup>erem* ‘scratch’, Gr. κείρω; \*g<sup>wh</sup> > ĵ in \*g<sup>wh</sup>ermo- > *ĵerm* ‘warm’, Gr. θερμός vs. \*g<sup>h</sup> > g- in *geĵk<sup>c</sup>* ‘glands’, Lith. *gėležuones*; and \*g > č in *čmem* (< \*čim-) ‘squeeze’, OCS *žьmъ* vs. \*g<sup>w</sup>enə<sub>2</sub> > *kin* ‘woman’, Gr. γυνή, OIr. *ben*. The attempt to consider palatalization the regular development with all velars and labiovelars alike (Kortlandt 1975, Martirosyan 2010: 711) is bought at the cost of excessive analogy.

### Sibilants

Word-initially the sibilant \*s is continued as either *h-* or zero without any clear conditioning, e.g. \*senos > *hin* ‘old’, Lith. *sėnas* vs. \*septm̥ > *ew<sup>c</sup>n* ‘seven’, Lat. *septem*, and



deleted intervocally, e.g. \*swesōr > \*k<sup>h</sup>eur > *k<sup>o</sup>oyr* ‘sister’. Two remaining problems are the word-final development, which may be either zero or *-k<sup>c</sup>* as in the plural marker, or most likely both (cf. the discussion of the nom. pl. ending) and the possibility of a special RUKI development that would connect Armenian with other satəm languages such as Indo-Iranian and Balto-Slavic. At least we find both *r* and *rš* from \*-rs-, perhaps as an indication of early dialectal variation, in the verb *r<sup>h</sup>airamim* vs. *r<sup>h</sup>aršamim* ‘wither, dry out’, Gr. τέρσομαι (cf. also Martirosyan (2010: 709), who hypothesizes an original accent distinction between the two variants), but it is an object of debate whether the high vowels \*i and \*u had a similar effect on a following \*s. If this is the case, as already suggested by Pedersen (1905: 228) and later argued by Winter (1975: 115ff.) and Olsen (1989a: 5–15), the result is *-r*, probably with an intermediate stage \*-ž, as, e.g., found in the aorist imperatives \*d<sup>h</sup>eh<sub>1</sub>-s > *dir* ‘put’ and \*deh<sub>3</sub>-s > *tur* ‘give’ from original injunctives, and *i*-stem verbal nouns like *dir* ‘position’ and *lir* ‘fullness’, which could represent the old *ti*-stems \*d<sup>h</sup>eh<sub>1</sub>tis (Av. *-dāiti-*, Goth. *dēds*, Lith. *dėtis*, OCS *-dětъ*) and \*pleh<sub>1</sub>tis (Ved. *prātī-*, Lat. *com-plētiō*) respectively.

### Liquids

The usual outcome of Indo-European \*r is *r*, with the variant *r̥* occurring before nasals and as a reflex of \*sr, cf. *berem* ‘carry’, Gr. φέρω vs. *beim* ‘burden’, Gr. φερνή ‘dowry’, and *aru* ‘stream’ < \*sruti- (Gr. ῥύσις ‘stream’, Ved. *sruti-* ‘road’). As is also the case in Greek and Anatolian, Armenian has a synchronic phonotactic constraint against initial *r-*, but it is a controversial issue whether this state of affairs should be projected back into the protolanguage. If this is the case we have to reconstruct the protoform of Ved. *rājas-*, Goth. *riqiz* ‘darkness’, Gr. ἔρεβος ‘the dark of the underworld’, Arm. *erek* ‘evening’ with an initial laryngeal, \*h<sub>1</sub>reg<sup>w</sup>os, and we would get a laryngeal-based “prothetic vowel” \*h<sub>1</sub>- > *e-* in Greek and Armenian. If, on the other hand, initial \*r- is accepted, we may reconstruct \*reg<sup>w</sup>os, which would automatically yield the same result (cf. also the section “Laryngeals”). The development of prothetic vowels where no laryngeal is involved is also found before \*r- in metathesized sequences (cf. “Consonant Clusters”) and in many Iranian loanwords with original initial *r-*, e.g. *aroyr* ‘brass’ from a North-West Iranian \*rauḏa-, originally ‘the red (material)’, *arag*, *erag* ‘quick’ vs. MParth. *rag* ‘quick’, Av. *rayu*.

The distribution between plain *l* and velarized *l̥* is partly predictable: the general rule stipulates *l* initially (e.g. *lizem* ‘lick’, Gr. λείχω) and intervocally (e.g. *gelum* ‘roll, turn’, Lat. *uoluō*) vs. *l̥* before consonants (e.g. *eln* ‘deer, hind’, Lith. *ėlnis* ‘deer’), but the original distribution is often blurred by analogy. In final position both variants are found, with a tendency toward *-l̥* in the oldest manuscripts (Meillet 1936: 47), thus e.g. *gayl̥* beside *gayl* ‘wolf’.

### Nasals

The nasals \*m and \*n are among the most stable phonemes in Armenian as well as in most other Indo-European languages. Nevertheless, the development in final position is open for discussion. The nasal is apparently lost in the acc. sg. of *o-* and *a-*stems as opposed to a merger of \*m and \*n into final *-n* in cases like \*ǵ<sup>h</sup>yōm > *jiwn* ‘snow’, Gr. χιών, or \*dōm > *tun* ‘house’. Similarly, a final vocalic \*-ṃ is continued as \*-n, e.g. acc. \*pod-ṃ > *otn* ‘foot’, Gr. πόδα. The original distribution can hardly be determined without a certain amount of analogy: Kortlandt (1984: 97) assumes that “original postvocalic \*-n

had apparently become a feature of the preceding vowel and subsequently been lost”; i.e. we have to posit an intermediate stage of nasalized vowels, and *jiwn* and *tun* must be analogical. Alternatively, one might suggest that nasalization only occurred with the open vowels \*o and \*a, in which case *jiwn* and *tun* would be regular while, on the other hand, the endless accusatives of animate *i*- and *u*-stems would have to be analogical.

### *Semivowels*

The fate of the semivowels \*y and \*w is still open for interpretation in several respects. While intervocalic \*-y- is lost at an early stage, e.g. \*treyes > *erek<sup>c</sup>* ‘three’, the development in word-initial position remains unclear due to the scarcity of diagnostic examples. Most likely, the regular development is *j*- as in *jur* ‘water’, which should probably be compared with Lith. *jūrės* ‘sea’. The initial *l*- of *leard* ‘liver’ and *luc* ‘yoke’ should not necessarily be considered counterevidence insofar as *leard* may constitute a contamination between \*yēk<sup>w</sup>rt̥ (Ved. *yákr̥t* etc.) and the stem of either OEng. *lifer* etc. < \*liparo- ‘fat’ or Hitt. *lišši-* ‘liver’, while *luc*, rather than being a direct continuation of \*(H)yugom (Ved. *yugám*, Gr. ζυγόν, Lat. *iugum*), may have been influenced by the verb *lucanem* ‘dissolve’. There are two reliable examples of initial *j*-, *ju* ‘egg’, SCr. *jaje*, and *jez* ‘you’ (acc./dat./loc. pl.), Ved. *yūyám*, but in both cases the protoform may have had two subsequent syllables in \*yV- as a disturbing factor. According to Martirosyan (2010: 706), who derives the interrogative/indefinite stem \*o- from the relative stem \*yo-, the regular continuation is zero. However, *o*- is preferably connected with the more appropriate stem \*k<sup>w</sup>o-, as \*k<sup>w</sup>- and \*p- seem to merge before \*-o-, cf. also \*k<sup>w</sup>olso- > *oln* ‘spine, backbone’, Lat. *collus*, *collum*.

In initial position \*w- is reflected as *g*-, no doubt via an intermediate stage \*gw- as also suggested by the early Georgian loanword *γwino* ‘wine’, Arm. *gini* < \*woynio- (Gr. οἶνος, Lat. *uīnum*), but the intervocalic development is more complicated (cf. the treatments by Eichner (1978), Olsen (1986) and Matzinger (1992)) with the reflexes -*g*-, -*w*- (-*v*-) and zero. Possibly -*g*- is the regular continuation before back vowels apart from -*u*-; -*w*/*v*- before front vowels and -*u*-; and zero between homorganic vowels: cf. e.g. *ōroganem/āroganem* ‘sprinkle’ from the root \*srew-, Ved. *srávati* ‘streams, flows’; *hoviw* ‘shepherd’ < \*h<sub>2</sub>owi-pah<sub>2</sub>-; gen./dat./loc. *aler* < \*-ewer- from *alewr* ‘flour’, Gr. ἄλειω < \*h<sub>2</sub>léh<sub>2</sub>w̥. An important detail has been added by Schindler (apud Matzinger 1992), who stipulated a regular development \*-wi- > -*wu*- whereby the otherwise puzzling transfer of original *i*-stems like *haw* ‘bird’, Lat. *avis*, to the *u*-declension has found an elegant solution.

### *Laryngeals*

The fate of the IE laryngeals is one of the most hotly debated issues of Armenian phonology, and while certain details seem relatively clear, there is hardly a single phonological conditioning where it has been possible to achieve full consensus. The following contexts are particularly noteworthy:

*Initial position before consonants (\*HC-)*: As is also the case of Greek and Phrygian, Armenian is characterized by so-called prothetic vowels deriving from laryngeals, with prominent examples like *astl* ‘star’, Gr. ἀστήρ, Hitt. *h(a)šterz* vs. Goth. *stairno*, Lat. *stella* < \*h<sub>2</sub>ster/l-; and *ayr* ‘man’, Gr. ἀνὴρ, Phryg. *avap* vs. Ved. *nár-*, Umbr. acc. pl. *nerf* < \*h<sub>2</sub>ner-. There is, however, no general agreement as to whether the prothetic vowel is always *a*-, irrespective of the quality of the laryngeal (thus, e.g., Klingenschmitt 1982:

105 and Olsen 1985), or Armenian shares the feature of “triple representation” with Greek as argued by, e.g., Winter (1965), Kortlandt (1987) and Beekes (1988), so that  $*h_1C- > *eC-$ ,  $*h_2C- > *aC-$ ,  $*h_3C- > *oC-$ . Since the adherents of “triple representation” at the same time favor a conditioned development of  $*o > a$  in such cases as *anicanem* ‘mock’ <  $*h_3neyd-$  (Gr. ὀνειδίζω, Goth. *ga-naitjan*), the evaluation depends on the correct interpretation of examples with  $*h_1-$ . This is further connected with the status of  $*r-$  in Indo-European: if  $*r-$  did occur in initial position in the protolanguage, *erek* ‘evening’, Gr. ἔρεβος may go back to  $*reg^wos$ , and examples such as *atamn* ‘tooth’, Aeol. pl. ἔδοντες, and *anun* ‘name’ with the possibly cognate Laconian names Ἐνομακρατίδας, Ἐνομμαντιάδας may be reconstructed with an initial  $*h_1-$ ; if not, the correct reconstruction of *erek* is  $*h_1reg^wos$ , which would then favor the idea of “triple representation”. This is, however, not what we find in other contexts. Prothetic vowels do not occur before  $w-$ , cf. *gelmn* ‘fleece, wool’ <  $*h_2wel_1m_ṛ-$ , Hitt. *hulana-*, and *goy* ‘(s)he is’ <  $*h_2wos-$ , Gr. ἄεσα νόκτα ‘spent the night’.

*Interconsonantal position* (\*CHC): When a laryngeal is “vocalized”, or more correctly develops a prop vowel interconsonantly, this vowel is always *a*, so at least in this position a “triple representation” matching the situation in Greek is excluded: thus *cnaw* ‘engendered’ <  $*gen_1to$ , Gr. ἐγένετο with  $*-a_1-$ ; *hayr* ‘father’ <  $*p_2tēr$ , Gr. πατήρ with  $*-a_2-$ ; *arawr* ‘plow’ <  $*h_2ar_3trom$ , Gr. ἄροτρον with  $*-a_3-$ . However, as is also the case in, e.g., Iranian, interconsonantal laryngeals are sometimes lost, as in *armn* ‘root’ <  $*h_2r_ṃno-$ , Gr. ὄρμενος or *dustr* ‘daughter’ <  $*d^hug_2tēr$ , Gr. θυγάτηρ. The conditioning of this diversity is not quite clear (cf. the discussion in Olsen 1999: 768f.), but at least vocalization seems to be regular in initial and final syllables.

*Position between vocalic sonant and stop/s* (\*RHC): Sequences of sonant + laryngeal in the position between consonants are reflected as either *aRa* or *aRaw* with a prop vowel on both sides of the sonant. Thus, *čanač'em* ‘know’ (for  $*canač'em$ ) <  $*g_ṃh_3-ske/o-$ , Gr. γινώσκω, as opposed to the corresponding derivative *canawt<sup>c</sup>* ‘known; knowing’ <  $*g_ṃh_3to/i-$ , where the development to *aRaw* may be restricted to peak position as originally suggested by Winter (1965: 110), cf. also *amač'em* ‘am ashamed’ beside *amawt<sup>c</sup>* ‘shame’ and *alac<sup>c</sup>em* ‘pray’ beside *alawt<sup>c</sup>k<sup>c</sup>* ‘prayer’.

*Initial position before vowels* (\*HV-): Apparently an initial consonantal laryngeal is sometimes reflected as *h-* in Armenian, while in other cases it simply seems to be lost. An attempt to account for this bewildering state of affairs has been made by Kortlandt (1980b) (followed by Beekes (1988: 76), Martirosyan (2010: 712f.) and Kloekhorst (2008: 75)), who further seeks to corroborate the theory with Anatolian evidence. According to Kortlandt’s analysis, the laryngeals  $*h_2$  and  $*h_3$  would be preserved in the protolanguage before an original *e*-vowel, while they would be lost before  $*o$ , i.e.  $*h_2e- > ha-$ ,  $*h_3e- > ho-$  vs.  $*h_2o$  and  $*h_3o > o$ . This distinction would still be reflected in Armenian, where an initial *h-* (if not from  $*s-$  or  $*p-$ ) is assumed to indicate an *e*-grade in the protoform. However attractive the idea of such a clear-cut regularity may be, the actual material is difficult to reconcile with the suggested distribution, as when, e.g., the word for ‘plow’, *arawr*, on account of the missing initial *h-* has to be analyzed as a zero-grade formation  $*h_2r_ṃtrom$  as opposed to the evidence for full grade in the other languages, or when the verb *acem* ‘bring, lead, conduct’ is considered a match of Lat. *gerō* <  $*h_2gese/o-$  rather than, as traditionally assumed, identical with Ved. *ājati*, Gr. ἄγω, Lat. *agō*.

*Position after i/u before a consonant* (i/uHC): Traditionally a laryngeal following the high vowels *i* and *u* is assumed to be lost with a compensatory lengthening that would leave no trace in Armenian, e.g. *jil* ‘sinew’ <  $*gwiHslah_2-$ , cf. Lith. *gýsla*, Lat. *fīlum*, or *c<sup>c</sup>uk<sup>c</sup>* ‘roof’ <  $*skuHto-$ , cf. Gr. σκῦτος. However, as also seems to be the case in Greek

and Tocharian (cf. Francis 1970: 276–84, Normier 1977: 182 and 1980: 273), it has been suggested (Olsen 1992, 1999: 770–73; cf. also Olsen 2009 on the conditioning in Greek) that \*h<sub>2</sub> and \*h<sub>3</sub> caused “laryngeal breaking”, i.e. \*ih<sub>2/3</sub> > \*ya/ea, \*uh<sub>2/3</sub> > \*wa/ea, thus, e.g., *erkar* ‘long’ < \*dwāro- < \*duh<sub>2</sub>ró- = Ved. *dūrā-*, Lat. *dūrus*, Dor. δῶρόν.

### Consonant clusters

A comparison between the syllable structure of the Indo-European protolanguage and that of Armenian makes it clear that very few original consonant clusters have survived in a historically transparent form. The treatment of several clusters remains unsettled, and at this place only a selection of the more well-established rules will be presented. A rich collection of potential examples may be found in Džaukjan 1982. Cf. also Godel 1975: 78–85.

### Clusters of two stops

The evidence for such clusters is rather scarce, but at least the word-internal development \*-pt- > -wʰ- is secured by *ewʰn* ‘seven’ < \*septm, analogically followed by *utʰ* ‘eight’ < \*optō (for \*okʷtō), while the regular outcome of a palatal stop (after *u*) + *t* is illustrated by *dustr* ‘daughter’ < \*dʰugə<sub>2</sub>tēr. Most likely, word-initial \*pt- is continued as *tʰ-*, cf. *tʰeli* ‘elm’: Gr. πτέλεα, *tʰer* ‘side’ < \*pter-, Gr. πτέρον ‘wing’.

### Clusters of resonant + stop

Voiceless stops are generally assumed to be voiced in the position after resonants, e.g. *bard* (*i-st.*) ‘pile’ < \*bʰrti-, Ved. *bhṛti-* ‘burden’; *hing* ‘five’, Ved. *pāñca*, Gr. πέντε. However, the cluster \*nt apparently has a double reflex, *nd* as in *ənderkʰ* ‘entrails’, Gr. ἔντερα, *dr-and* ‘doorposts’, Lat. *antae*, vs. *n* in, e.g., the 3 pl. ending *-(V)n* < \*-(V)nti, or decadic numbers in *-sun*, e.g. *eresun* ‘30’, Gr. τριάκοντα. The conditioning is differently interpreted: Olsen (1989b: 227) stipulates voicing after a stressed vowel, while Clackson (1994: 56) assumes *-nd-* in internal position vs. *-n* word-finally after the loss of end-syllables. Somewhat similarly, the intervocalic clusters \*mp and \*mbʰ appear to have two outcomes, *mb* and *m*, e.g. \*mp > *mb* in *tʰamb* ‘(pack) saddle, ham’, ONor. *þomb* ‘paunch; bowstring’, Lith. *tūpa* ‘sinew’ vs. \*mp > *m* in *amokcʰem* ‘soften; mix; cook, boil together’, Skr. *saṃ-pacati* (cf. Dumézil 1938 and Olsen 2005 on a possible regular distribution).

### Clusters of resonant + s/s + resonant

For resonant + s/s + resonant we observe the following rules:

- |                |  |
|----------------|--|
| *Ns > s        | e.g. *mēmso- > <i>mis</i> ‘flesh, meat’, Ved. <i>māmsá-</i> , Goth. <i>mimz</i>  |
| *ls > l        | e.g. <i>al</i> ‘salt’, Gr. ἄλς   |
| *rs > r̥ or rš | e.g. <i>tʰarāmim/tʰaršamim</i> ‘wither dry out’, Gr. τέρσομαι (cf. “Sibilants”)  |
| *sN > N        | e.g. <i>gin</i> ‘prize’, Ved. <i>vasná-</i> , Lat. <i>uēnum</i> ; <i>nu</i> ‘daughter-in-law’ < *snu-sos, Gr. νυός         |
| *sr > r̥       | e.g. <i>kʰer̥</i> < *swesros, gen. of <i>kʰoyr</i> < *swesōr ‘daughter’; <i>aru</i> ‘stream’ < *sruto-, Ved. <i>srutá-</i> |

## Clusters of stop + resonant

In initial position a voiceless stop is lost at least before a liquid with further development of a prothetic vowel before what would otherwise be initial *r-*, thus \*pleh<sub>1</sub>tos > *li* ‘full’, Lat. *-plētus*; \*k<sup>w</sup>lutos > *lu* ‘heard’, Gr. κλυτός, Ved. *śrutá-*, Lat. *-clutus*; \*treyes > *erek* ‘three’, Ved. *tráyah*, Gr. τρεῖς, Lat. *trēs*; \*preysg<sup>w</sup>us > *erēc* (*u-st.*) ‘elder’, Cret. πρεῖσγος. Word-internally, we generally find a lenited version of the stop:

*-tr- > -wr(-),	e.g. <i>arawr</i> ‘plow’, Gr. ἄροτρον, Lat. <i>arātrum</i> ; <i>hawr</i> < *pə <sub>2</sub> trós, gen. of <i>hayr</i> ‘father’
*-kr- or *-kr- > -wr(-)	e.g. <i>mawruk</i> ‘beard’, Ved. <i>śmáśru-</i> ‘beard’, Lith. <i>smākras</i> ‘chin’
*-tl- > -wl(-)/-wl(-)	e.g. <i>ul</i> (for *uwl; <i>o-st.</i> ) ‘kid’ < *putlos, Osc. <i>puclō-</i> ‘son’; agent noun suffix <i>-awl</i> , e.g. <i>cnawl</i> ‘parent’ < *-VtlV- (cf. “Nominal Word Formation”)
*-pn- > -wn(-)	e.g. <i>tawn</i> ( <i>i-st.</i> ) < *dapni- ‘feast’, Lat. <i>daps</i> ‘sacrificial meal’, ONor. <i>tafn</i> ‘victim’

## Metathesis

Metathesis regularly occurs in clusters of media or voiced aspirate + liquid, triggering the development of a prothetic vowel word-initially, cf. the following examples:

*dr	>	<i>rt</i>	*swidro- > <i>k’irtn</i> ‘sweat’, Latv. <i>sviēdri</i>
*g <sup>w</sup> r-	>	<i>Vrk-</i>	*g <sup>w</sup> rah <sub>2</sub> wŋ- > <i>erkan</i> ‘grindstone’, Ved. <i>grāvan-</i>
*b <sup>h</sup> r	>	<i>rb</i>	*kub <sup>h</sup> ro- > <i>surb</i> ‘pure, holy’, Skr. <i>śubhra-</i>
*g <sup>h</sup> r	>	<i>rj</i>	*meġ <sup>h</sup> ri > <i>merj</i> ‘near’, Gr. μέχρη
*g <sup>(w)h</sup> r-	>	<i>Vrg-</i>	<i>argand</i> ‘womb’, Slav. <i>grōdb</i> ‘breast’

In case of the structure \*Vr – r there is regular dissimilation to (V)l – r, thus, e.g., \*b<sup>h</sup>rātēr > *elbayr* ‘brother’, \*b<sup>h</sup>rēwŋ > *albewr* ‘fountain’, Gr. φρέαρ.

A small group of exceptions includes *srunk*<sup>c</sup> ‘leg, shank’ < \*krūsni- or the like (Lat. *crūs* ‘shank’, Lith. *šlaur̃nis* ‘hip, thigh’), not \*Vrunk<sup>c</sup>; *krunk* ‘crane’, Lat. *grūs*, not \*Vrkunk; and possibly *glux* ‘head’, Lith. *galvā*, etc., if this is from \*g<sup>h</sup>luH- rather than the morphologically unlikely \*g<sup>h</sup>ōlū-. These examples share the features of an initial velar or palatal and *u*-vocalism, so perhaps a sequence \*KRu- would develop a prop vowel to \*KuRu- whence, with loss of the unaccented *-u-*, *KRu-*.

## Clusters of two resonants

In a few cases we have evidence for a development \*ln- > -l-, e.g. *ʳohum* ‘let’ < \*tol-nu-, Lat. *tollō* (cf. Klingenschmitt 1982: 243f.), while it is more difficult to determine the conditioned distribution of *wn* beside *mn* from the sequence \*mn (cf. the discussion in Olsen 1999: 792f.).

Clusters of *s* + stop/stop + *s*

The cluster \*st is preserved initially as well as in original intervocalic position, thus *stipem* ‘press’, Gr. στειβω; *sterj* ‘barren’, Gr. στειρα; similarly \*zd in *ost* ‘branch’ < \*(H)o-zdo-

Gr. ὄζος. The regular reflex of \*sp- is more uncertain: *sp-* as in *sparnam* ‘threaten’, Lat *spernō*, or *p<sup>c</sup>-* as in *p<sup>c</sup>oyr<sup>c</sup>* ‘zeal’, Gr. σπουδή. An even more complicated matter is the fate of *sK/Ks*-clusters. According to Godel (1975: 80), *k* and \**k* are neutralized in connection with \**s*, and, indeed, we find quite a number of sound etymologies including such a cluster with a common reflex *c<sup>c</sup>*, e.g. \*-*ks* in *vec<sup>c</sup>* ‘six’ < \**suweks*, \*-*sg<sup>w</sup>-* in *erēc<sup>c</sup>* ‘elder’ < \**preysg<sup>w</sup>us*, Cret. πρεῖσγυς, \**sk-* in *c<sup>c</sup>urt* ‘cold’, OCS *sěverъ* ‘north’, and \**sk-* in *c<sup>c</sup>ayfem* and *c<sup>c</sup>aytem* ‘bubble, sparkle’, Lat. *scateō*, Lith. *skaidrūs*; similarly, the original three-consonant clusters \*-*s-sk-* in *ayc<sup>c</sup>* ‘visit, search’ < \**h<sub>2</sub>ays-sk-* and \*-*k-sk-* in *har<sup>c</sup>anem* ‘ask’ < \**ṗrk-sk-*, Ved. *ṗṛcchāti*. Another potential development is *-č-*, which may be found in \*-*sk-*verbs of the type *čanačem* ‘know’ < \*(*gi-*)ḡnh<sub>3</sub>-*ske-*, Gr. γιγνώσκω, *atačem* ‘pray’ < \*(*si-*)sḷh<sub>2</sub>-*ske-*, Gr. ἰάσκεσθαι, but here it is a matter of dispute whether the Armenian verbs represent a faithful continuation of the Indo-European pattern (thus, e.g., Klingenschmitt 1982: 84) or an extended suffix \*-*sky-*. Finally, there exists a small group where an initial *š-* apparently goes back to \**sk-* (or \**skh-*, cf. Martirosyan 2010: 516): *šel* ‘oblique’ < \**skel-no-*, Gr. σκέλλος ‘bandy-legged’, Alb. *çalë* ‘lame’, Lith. *kušnas* ‘heel’; *šil* ‘cockeyed’, OHG *skēlah*; *šert* ‘splinter’, Lith. *skedervà*; *šertem* ‘scratch’ < \**skerd-*, Lith. *skėrdžiu*. In all these cases, the clusters occur at the same time initially and before front vowels.

### Y-clusters

The correct interpretation of the fate of *y*-clusters is severely hampered by the scarcity of reliable examples. Thus, the regular continuation of \**ty* and \**dy* is assumed to be either the “shibilants” *č<sup>c</sup>* and *č* or the affricates *c<sup>c</sup>* and *c*, while the development of \**d<sup>h</sup>y* > *j* is quite uncontroversial, cf. *měj* ‘middle’ < \**med<sup>h</sup>yo-* (with \**e* > *ē* before a palatal, cf. also \*(*h*)ekwos > *ěš* ‘donkey’), Ved. *mádhya-*, Lat. *medius*. Truly, the assumed continuations *č<sup>c</sup>* and *č* would mean a neat parallelism between the three series, but the best example of \**ty* > *č<sup>c</sup>* is the verb *kočem* ‘call’, allegedly < \**g<sup>w</sup>otye-* and thus a cognate of Goth. *qīpan* ‘say, speak’, and this may have been influenced by the almost synonymous *gočem* < \**wok<sup>w</sup>ye-* ‘shout, cry, call out’. For \**dy* > *č* the phonetic details of the alleged connection between *oročam*, *-em* ‘chew, ruminate’ and Lat. *rōdō*, Ved. *rādati* ‘gnaw’ (Greppin 1993: 19) are unclear. According to an alternative theory (Olsen 1993), \**ty* is continued as the affricate *c<sup>c</sup>* in, e.g., *k<sup>c</sup>ec<sup>c</sup>* ‘isolated’ < \**swetyo-*, Lith. *svėčias* ‘stranger, guest’ (from the reflexive stem \**swe-*) and the verb *ancanem* ‘pass, cross’, Gr. ἀντίομαι ‘stand opposite’, ἀντιάω ‘come toward’ (cf. also under “Verbs” on the subjunctive). The existence of parallel examples of \**dy* > *c* (cf. Godel 1975: 82) largely depends on the interpretation of verbs such as *hecanim* ‘ride’, Gr. ἔζομαι ‘sit’, *anicanem* ‘mock’, Gr. ὀνειδίζω, Goth. *ganaitjan* as the continuations of *y*-presents rather than *s*-aorists. The development of \**k<sup>w</sup>y* > *č<sup>c</sup>* is safely attested in *ač<sup>c</sup>k<sup>c</sup>* ‘eyes’, Gr. ὄσσε, OCS *oči*, while \**ky* most likely yields *c<sup>c</sup>*, cf. *lucanem* ‘lighten’ < \**lōwk<sup>w</sup>ye-* (\**k* > \**k* after *u*-diphthong; Klingenschmitt 1982: 194). Intervocalic \*-*sy-* is continued as *-y-*, thus the *o*-stem gen. sg. ending \*-*osyo* > *-oy*, and in the position after a sonant, \*-*y-* is reflected as *-j(-)*, e.g. *anurj* ‘dream’, Gr. ὄνειρος.

### W-clusters

Some of the most bizarre phonetic developments in Armenian are found with *w*-clusters, though the end results are less surprising when we consider both the strengthening of \**w* > *g* and the general tendency toward metatheses. Thus, \**sw* and \**tw* are reflected as *k<sup>c</sup>* in, e.g., \**swesōr* > *k<sup>c</sup>oyr* ‘sister’, Ved. *svásar-* and \**two-* > *k<sup>c</sup>o* ‘your’, Ved. *tva-*. As famously discovered by Meillet (1924) \**dw* is continued as (V)*rk* in examples such as *erku* ‘two’,

Lat. *duo* and *erkar* ‘long’, Gr.  $\delta\eta\rho\acute{o}\varsigma$ . This remarkable sound law can hardly be dismissed, though it has often been doubted and its implications widely debated, cf. Klingenschmitt 1982: 225 and 238f., Vennemann 1986, Kortlandt 1989 and Viredaz 2003. On this background one might expect a parallel development  $*d^hw > rg$ , and in fact this is what Klingenschmitt (1982: 238f.) assumes for the verb *argelum* ‘hinder, keep back’, which he compares with the synonymous ONor. *duelia*, OHG *twellen*. The regular continuation of the cluster  $*kw$  is  $\acute{s}$ , cf. *šun* ‘dog’, Ved. *śvā́*; *ēš* ‘donkey’, Ved. *ásva-*.

## MORPHOLOGY

The characteristic features of the Indo-European morphological system are only partly preserved in Armenian, with a number of restructurings, simplifications and losses in both the noun and the verb. Several endings and morphological markers have been rendered opaque by regular sound change, including the loss of final syllables, and to the extent the corresponding categories are maintained, we observe a series of countermoves against the regular phonetic development such as the construction of suffix clusters and the addition of clarifying postpositions or adverbs. However, except for the numerous Iranian noun suffixes, the morphological elements as such generally seem to have belonged to the protolanguage.

### Nouns

The system of noun inflection in Classical Armenian has undergone significant changes: the category of grammatical gender has been given up altogether, the dual is lost, and the case system, while preserving a distinction between seven cases in the pronouns (the inherited eight minus a separate vocative), has been subject to several syncretisms, so that a paradigm will usually only distinguish between four different forms in the singular and four in the plural. On the other hand, an archaic ablaut pattern has been preserved to some extent, particularly in the inflection of the *n*-stems and a few inherited kinship terms.

#### Vocalic stem classes

The inflection of the four vocalic stem classes, *o*-, *a*-, *i*- and *u*-stems, is almost identical. However, the *a*- and *i*-stems merge in the gen./dat./loc. sg., and the *o*-stems exhibit some aberrations in the oblique cases of the singular.

#### Vocalic stem classes:

	<i>o</i> -stems	<i>a</i> -stems	<i>i</i> -stems	<i>u</i> -stems
	<i>ēš</i> ‘donkey’	<i>am</i> ‘year’	<i>ban</i> ‘word’	<i>zard</i> ‘ornament’
sg. nom./acc.	<i>ēš</i>	<i>am</i>	<i>ban</i>	<i>zard</i>
gen./dat.	<i>iš-oy</i>	<i>am-i</i>	<i>ban-i</i>	<i>zard-u</i>
abl.	<i>iš-oy</i>	<i>am-ē</i>	<i>ban-ē</i>	<i>zard-ē</i>
instr.	<i>iš-ov</i>	<i>am-aw</i>	<i>ban-iw</i>	<i>zard-u</i>
loc.	<i>ēš</i>	<i>am-i</i>	<i>ban-i</i>	<i>zard-u</i>
pl. nom.	<i>ēš-k<sup>c</sup></i>	<i>am-k<sup>c</sup></i>	<i>ban-k<sup>c</sup></i>	<i>zard-k<sup>c</sup></i>
gen./dat./abl.	<i>iš-oc<sup>c</sup></i>	<i>am-ac<sup>c</sup></i>	<i>ban-ic<sup>c</sup></i>	<i>zard-uc<sup>c</sup></i>
acc./loc.	<i>ēš-s</i>	<i>am-s</i>	<i>ban-s</i>	<i>zard-s</i>
instr.	<i>iš-ovk<sup>c</sup></i>	<i>am-awk<sup>c</sup></i>	<i>ban-iwk<sup>c</sup></i>	<i>zard-uk<sup>c</sup></i>

A subtype of the *a*-stems with gen./dat./loc. sg. *-ay* is practically restricted to proper names, e.g. *Vardan*, *-ay* (Weitenberg 1989).

The numerous case syncretisms are naturally explained by the fact that final syllables would generally be lost. Thus, the endingless forms of the nom./acc. sg. are probably the regular continuations of *\*-os*, *\*-om* (*o*-st.), *\*-ah<sub>2</sub>*, *\*-ah<sub>2</sub>m* (*a*-st.), *\*-is*, *\*-i* (*i*-st.) and *\*-us*, *\*-u* (*u*-st.); the missing ending of the old animate accusative of *i*- and *u*-stems, *\*-im* and *\*-um*, may be either regular or analogical. In the *o*-stems the disyllabic ending of the gen. sg. *\*-osyo* > *-oy* (cf. Skr. *-asya*) was extended to the dative and the ablative, where the inherited ending *\*-ōd* or *\*-ād* would undergo apocope, and the endings of the instr. sg. are unproblematically derived from the stem vowel + *\*-b<sup>h</sup>i* (*-u* representing *\*-uw*), but otherwise most of the inherited endings would have been lost by apocope. Therefore, the loc. sg. of the *o*-stems is generally endingless (IE *\*-oy* > *-Ø*), and the gen./dat. of the *a*- and *i*-stems in *-i* and that of the *u*-stems in *-u* must go back to disyllabic forms *\*-i-os*, *\*-u-os*, etc. The locative in *-i*, also occasionally found with the *o*-stems, most likely has its origin in a postposition *\*-h<sub>1</sub>en*. Similarly, the ablative in *\*-ē* may be explained as an original postposition *\*-eti*, cf. Ved. *āti* ‘over, beyond’. If intervocalic *\*-t-* is continued as *\*-w-* before a back vowel, as seems to be the case, the traditional derivation from *\*-etos* corresponding to Ved. *mukhatāḥ* ‘from the mouth’ etc. (Meillet 1936: 73) is phonetically impossible.

In the plural the most troublesome ending is the nominative in *-k<sup>c</sup>* where one expects a final *\*-s*. Klingenschmitt (1982: 23f.) explains *-k<sup>c</sup>* in the nom. pl. beside the zero ending in the nom. sg. as sandhi variants: when in the *a*-stems the singular *\*-ah<sub>2</sub>* > *-Ø* was opposed to the plural *\*-ah<sub>2</sub>(e)s* > *\*-Ø/-k<sup>c</sup>*, *-k<sup>c</sup>* was perceived as the characteristic plural marker, and in the course of time *-k<sup>c</sup>* was generalized in all plural forms, nominal as well as verbal, thus also instr. pl. *-ovk<sup>c</sup>* etc., 1 pl. verbal ending *-emk<sup>c</sup>* etc., pronouns such as *mek<sup>c</sup>* ‘we’, *duk<sup>c</sup>* ‘you’, and, perhaps most significantly, *erek<sup>c</sup>* ‘three’ < *\*treyes* and *č<sup>c</sup>ork<sup>c</sup>* ‘four’ < *\*k<sup>w</sup>et(w)ores*, the only numerals ending in *\*-s*. Alternatively, Kortlandt (1984: 89f.) assumes that *-k<sup>c</sup>* is the only regular reflex of postvocalic *\*-s*. In that case, however, the nominative must have been analogically replaced by the accusative, and it would be difficult to understand the general merger of masculine *o*-stems and neuter *s*-stems.

The *-s* of the acc./pl. is the regular outcome of the m./f. ending *-Vns* in the vocalic classes, while the same ending in the loc. pl. may have been extended from the *n*-stems, i.e. *\*(V)n-su* > *-s* with analogical restitution of the *n* in the *n*-stems themselves, e.g. *harsn* ‘bride’, acc./loc. pl. *harsuns* for regular *\*harsus*. Another possible source is the *s*-stems, where a geminate *\*-s-su* would probably yield *-s* as in *\*h<sub>1</sub>es-si* > *es* ‘you are’, while *\*-s-* of the vocalic stems would be lost in intervocalic position without a trace. The origin of the plural oblique ending *-(V)c<sup>c</sup>* remains a bone of contention. An attractive solution has been offered by Seldeslachts (1991: 261), who suggests a derivation from the original adjective suffix *\*-isko-* > *-ic<sup>c</sup>-*, later analogically adapted to the other stem classes as *-oc<sup>c</sup>-*, *-ac<sup>c</sup>-* etc. Finally, the instr. pl. is formed by adding the plural marker *-k<sup>c</sup>* to the instr. sg. Cf. also Matzinger 2005 for a more detailed discussion of the case endings.

In some cases the vocalic stems are direct continuations of the inherited patterns, thus *ēš* (*o*) ‘donkey’ < *\*(h<sub>1</sub>)ekwo-* ‘horse’, Lat. *equus* etc.; *am* (*a*) ‘year’ < *\*sm<sup>h</sup>(h)ah<sub>2</sub>-*, Ved. *sāmā-* ‘season’; *ban* (*i*) ‘word’ < *\*b<sup>h</sup>ah<sub>2</sub>-ni-*, ONor. *bón*; *z-ard* ‘ornament’ (*u*) < *\*-h<sub>2</sub>r<sub>1</sub>tu-*, Ved. *r<sub>1</sub>tū-* ‘rule, order’, Lat. *artus* ‘limb’, but they have also been fed by moribund categories and by loanwords. Thus, e.g., neuter *s*-stems like *jet* ‘tail’, Av. *-zadah-*, have been included in the *o*-stems, and root nouns like *ayc* ‘goat’, Gr. *αἴξ*, in the *i*-stems, presumably under the influence of the oblique plural *\*-isko-* > *-ic<sup>c</sup>-*. Compare also the old root noun *otn* ‘foot’ where the acc. sg. *\*pod-m̄* explains the singular inflection as an *n*-stem, while the plural, *otk<sup>c</sup>*, *otic<sup>c</sup>*, is treated as an *i*-stem. A number of



*u*-stems such as *haw* ‘bird’, Lat. *avis*, go back to older *i*-stems as a consequence of the regular development of *\*-wu-* > *-wi-*.

The oldest layer of Iranian loanwords seems to have preserved the inflectional class they had in Old Iranian, as originally assumed by Meillet (differently, Schmitt 1983: 100), e.g. the *i*-stem *uxt* ‘oath’, Av. *uxti-* ‘word, speech’; the *u*-stem *xrat* ‘advice’, Av. *xratu-*; or, not least, the numerous *a*-stem loanwords and loan suffixes going back to Indo-European thematic stems.

### *Alternating stem classes*

Except for the regular syncope and vowel reduction in unaccented syllables, the vocalic classes display no stem alternation. With the consonantal subtypes, i.e. the *n-*, *r-* and *l-*stems, on the other hand, a distinction is made between at least the stem of the nom./acc. sg. and that of the oblique forms.

### *N-stems*

From a historical point of view, the Armenian *n*-stems have several sources: some, like *gairn* ‘lamb’, Gr. ἀρήν, go back to various types of Indo-European *n*-stems, including *-men*-stems, Hoffmann formations and *-nt*-stems; others have been adapted from original root nouns, e.g. *otn* ‘foot’ from acc. sg. *\*podm̥*, and still others derive from stems in *\*-no-/-na-* such as *t<sup>o</sup>rn* ‘grandchild’ < *\*tornō-*, Lith. *tařnas* ‘servant, farm boy’ and *bern* ‘burden’ < *\*b<sup>h</sup>ernah<sub>2</sub>-*, Gr. φερνή ‘dowry’. As an archaic feature of the *n*-stems the paradigms display three or even four stem alternants: *n/in/an*, *n/un/an* and *n/in/an/un*, where *-in-*, *-un-* and *-an-* in principle represent the *e-*, *o-* and zero grade of the suffix *\*-en-*, *\*-on-*, *\*-n̥-*, so that it is still possible to distinguish between various inherited ablaut types. Thus, e.g., *anjn* ‘person; self’, gen. *anjin*, nom. pl. *anjink<sup>c</sup>*, continues a hysterodynamic paradigm corresponding to ONor. *angi* ‘smell’, while the compound *mi-anjn* ‘one-person’, i.e. ‘monk’ with gen. *mi-anjun*, nom. pl. *mi-anjunk<sup>c</sup>*, points to a compositional *o*-grade in the suffix with a similar allomorphy as Gr. πατήρ, acc. πατέρα ‘father’ vs. εὐπάτωρ, acc. εὐπάτορα. An example of an alternating (proterodynamic) paradigm is *harsn* ‘bride’ < *\*prkōn* (root *\*prek-* ‘ask’; cf. Hamp 1988).

### The *n*-stem *harsn* ‘bride’

sg. nom./acc.	<i>harsn</i>	< <i>*-ōn</i>	pl. nom.	<i>harsunk<sup>c</sup></i>	< <i>*-on-</i>
gen./dat./loc.	<i>harsin</i>	< <i>*-en-</i>	acc./loc.	<i>harsuns</i>	< <i>*-on-</i>
abl.	<i>harsnē</i>	< <i>*-en-</i>	gen./dat./abl.	<i>harsanc<sup>c</sup></i>	< <i>*-n̥-</i>
instr.	<i>harsamb</i>	< <i>*-n̥-</i>	instr.	<i>harsambk<sup>c</sup></i>	< <i>*-n̥-</i>

An extremely widespread subtype consists of nouns in *-iwn*, in particular abstract nouns in *-ut<sup>c</sup>iwn*, gen. *-u-t<sup>c</sup>ean*, e.g. *cerut<sup>c</sup>iwn* ‘old age’ from *cer* ‘old’. Finally, the monosyllabic *tun* ‘house’, Gr. δῶ, and *řun* ‘dog’, Ved. řvá, Gr. κύων, have the oblique forms *tan* and *řan*.

*N*-stem plurals are also found in some nouns denoting persons or animals, e.g. *nu* (gen. *nuoy*) < *\*snuso-* ‘daughter-in-law’ with the plural *nuank<sup>c</sup>*, *erēc<sup>c</sup>* ‘elder’ < *\*preysg<sup>w</sup>us* with the plural *erīc<sup>c</sup>unk<sup>c</sup>*. Here the *n*-extension probably corresponds to the individualizing *n*-stem suffix of Germanic.

**R-stems**

On the whole, the *r*-inflection is a relatively rare type, but, still, it contains some of the most essential elements of the vocabulary, the basic kinship terms. We may distinguish between three subtypes:

- 1) “regular” *r*-stems in nom./acc. sg. *-r*, gen./dat./loc. *-er*, e.g. *dustr*, *dster* ‘daughter’ < nom. \*d<sup>h</sup>ugə<sub>2</sub>tēr, with generalized suffix *-ter-* in the rest of the paradigm; similarly loanwords such as *kaysr*, *kayser* ‘emperor’.
- 2) a small number of \*-wer/-wen-heteroclitics, e.g. *alewr* (later *aliwr*) ‘flour’ < \*h<sub>2</sub>lēh<sub>1</sub>w<sub>1</sub>, Gr. ἄλειωρ, gen./dat./loc. *aler* (Eichner 1978).
- 3) the kinship terms *hayr* ‘father’, *mayr* ‘mother’ and *elbayr* ‘brother’ with preservation of the inherited ablaut pattern, here exemplified by the paradigm of *hayr*.

**Paradigm of *hayr* ‘father’**

sg. nom./acc.	<i>hayr</i>	< *pə <sub>2</sub> -tēr	pl. nom.	<i>hark<sup>c</sup></i>	< *pə <sub>2</sub> -tér-es
gen./dat./loc.	<i>hawr</i>	< *pə <sub>2</sub> -tr-ós etc.	acc./loc.	<i>hars</i>	< *pə <sub>2</sub> -tér- + -s
abl.	<i>hawrē</i>	< *pə <sub>2</sub> -tr- + -ē	gen./dat./abl.	<i>harc<sup>c</sup></i>	< *pə <sub>2</sub> -tr̥-sko-
instr.	<i>harb</i>	< *pə <sub>2</sub> -tr̥-b <sup>h</sup> i	instr.	<i>harbk<sup>c</sup></i>	< *pə <sub>2</sub> -tr̥-b <sup>h</sup> is

The paradigm of *k<sup>c</sup>oyr* ‘sister’ is quite isolated and highly archaic: nom./acc. sg. *k<sup>c</sup>oyr* < \*swesōr; gen./dat./loc. *k<sup>c</sup>er* < \*swesros etc.; instr. *k<sup>c</sup>erb* < \*swesrb<sup>h</sup>i, analogical for \*k<sup>c</sup>arb < \*k<sup>c</sup>earb; nom. pl. *k<sup>c</sup>ork<sup>c</sup>* < \*swesores.

**L-stems**

Besides a few secondarily adapted loanwords like *skutl*, gen. *sktel* ‘plate’, Lat. *scutella*, the only member of this stem class is *astl*, gen. *astel* ‘star’. The reason why this word is an *l*-stem rather than an *r*-stem like, e.g., Gr. ἄστῆρ is unclear: perhaps analogical influence from the word for ‘Sun’, \*sah<sub>2</sub>wōl, which is, however, not transmitted in Armenian, or perhaps even a trace of an old *r/l*-alternation as suggested by Olsen (2010: 81).

**Heteroclitics and irregular paradigms**

A number of *u*-stem adjectives exhibit a peculiar heteroclitic pattern with *-r* in the nom./acc. sg. and *n*-stem inflection in the plural, e.g. *barjr* ‘high’, gen./dat./loc. *barju*, nom. pl. *barjunk<sup>c</sup>* from the root \*b<sup>h</sup>erǵ<sup>h</sup>-, where the *u*-stem has a parallel in Hitt. *parkuš* and the *n*-stem may go back to the same *nt*-stem as Ved. *brhánt-*, Av. *bərəzant-*. The origin of the *-r* of the nom./acc. is more doubtful. It hardly reflects the competing adjective suffix \*-ro- as in Toch. A. *pärkär*, B. *pärkare* since *u*-adjectives in general seem to be replaced by *ro*-adjectives in Tocharian. Moreover, a nom./acc. in *-r* also occurs in a few original *u*-stem neuters, *artawsr*, pl. *artasuk<sup>c</sup>* ‘tear’ < \*draku, cf. Gr. δάκρυ; *cunr* ‘knee’, Gr. γόνυ; and *metr* ‘honey’, which is usually considered a contamination between \*melit ‘honey’ (Hitt. *melit* etc.) and the *u*-stem \*med<sup>h</sup>u, Ved. *mādhu* ‘honey, mead’, while it is missing in masculine *u*-stems such as *z-gest* ‘clothing, garment’ < \*westu-. On the whole, the best option may well be an idea by Alexis Manaster Ramer (p.c.), who stipulates a regular sound change of final *-u* > *-r*, which would imply that the neuter gender was generalized in the adjectives.

A modest number of nouns cannot be fitted into the regular inflectional types, thus, e.g., *akn* ‘eye’, gen. *akan*, pl. *ač<sup>c</sup>k<sup>c</sup>*, cf. Gr. ὄσσε; *unkn* ‘ear’ with the plural *akanj<sup>c</sup>*; *ayr* ‘man’ < \*h<sub>2</sub>nér, Gr. ἀνὴρ, gen. *arn*, metathesized from \*h<sub>2</sub>nṛós, Gr. ἀνδρός; *kin* ‘woman’ < \*g<sup>w</sup>enə<sub>2</sub> with the enigmatic oblique form *knoj* and the plural *kanayk<sup>c</sup>* with a remarkable similarity to Gr. γυναικες.

### Gradation

The Armenian adjective has no obligatory gradation (cf. Jensen 1959: 68f.). For the expression of the comparative, the positive is generally used either alone or with *k<sup>c</sup>an* (= Lat. *quam*) or *k<sup>c</sup>an z-*, cf. e.g. Matt. 18.6: *law ē nma* ‘it is better for him’ or Mark 14.5: *p<sup>c</sup>ok<sup>c</sup>r ē k<sup>c</sup>an zamenayn sermanis* ‘it is smaller than all seeds’. However, the original intensive in *-agoyñ*, an Iranian loan suffix, also serves as a comparative. For the expression of the superlative function there are several possibilities: the simple positive, e.g. Matt. 5.19: *p<sup>c</sup>ok<sup>c</sup>r koč<sup>c</sup>esc<sup>c</sup>i* ‘he shall be called the smallest’; reduplication, e.g. *mecamec* ‘the largest’ (or more commonly ‘very large’); or compounds with *amena-* ‘of all’, e.g. *amenasurb* ‘the holiest’ (also ‘very holy’).

### Pronouns

The pronominal system of Armenian consists of the following categories: the personal pronoun; the reflexive, reciprocal and possessive pronouns; the interrogative, indefinite and relative pronouns; and the demonstrative pronouns. Only the pronouns have a distinction between the nominative and accusative and between the genitive and dative of the singular.

#### The personal pronouns

<b>Singular</b>	First person	Second person
nom.	<i>es</i>	<i>du</i>
acc./loc.	<i>is</i>	<i>k<sup>c</sup>ez</i>
gen.	<i>im</i>	<i>k<sup>c</sup>o</i>
dat.	<i>inj</i>	<i>k<sup>c</sup>ez</i>
abl.	<i>inēñ, injēñ</i>	<i>k<sup>c</sup>ēñ, k<sup>c</sup>ezēñ</i>
instr.	<i>inew</i>	<i>k<sup>c</sup>ew</i>
<b>Plural</b>		
nom.	<i>mek<sup>c</sup></i>	<i>duk<sup>c</sup></i>
acc.	<i>mez</i>	<i>jez</i>
gen.	<i>mer</i>	<i>jer</i>
dat./loc.	<i>mez</i>	<i>jez</i>
abl.	<i>mēñj</i>	<i>jēñj</i>
instr.	<i>mewk<sup>c</sup></i>	<i>jewk<sup>c</sup></i>

1 sg.: the nominative *es* seems to be a sandhi variant for \*ec < \*(h<sub>1</sub>)eğō, while the gen. *im*, from an original possessive pronoun \*(h<sub>1</sub>)emo-, with assimilation to *in-* from the acc./loc. *is* < \*im-s, dat. *inj*, apparently provided the stem of the remaining case forms. The dative contains the same element \*-ğ<sup>b</sup>i as Lat. *mihī*.

2 sg.: nom. sg. *du* for expected \*t<sup>u</sup> < \*t<sup>ũ</sup> has the same (enclitic) development of \*t- as the demonstrative *d-* < \*to-. Here again, gen. *k<sup>c</sup>o* < \*two- (/twe-) is the old

possessive pronoun whose stem has spread to the rest of the paradigm. The dat. *k<sup>c</sup>ez* < \**twe-ǵ<sup>hi</sup>* must be a substitution for \**-b<sup>hi</sup>*, cf. Lat. *tibī*, under the influence of the 1 sg.

1 pl. *mek<sup>c</sup>* goes back to either \**mes* like Lith. *mēs* in analogy with the verbal ending or \**meyes* for \**weyes* with initial \**w-*, cf. Skr. *vayām*, Lat. *uōs* etc. The acc./dat./loc. *mez* has probably taken over the element \**-ǵ<sup>hi</sup>* from the singular, and the gen. *mer* contains the suffix \**-(t)ero-*, cf. Gr. ἡμέτερος.

2 pl. *duk<sup>c</sup>* for \**yūs* > \**juk<sup>c</sup>* (?) is influenced by the singular *du*.

### *The reflexive, reciprocal and possessive pronouns*

The reflexive pronoun *iwr* is common for the genitive, dative and locative; there also exists an abl. *iwr<sup>mē</sup>*, an instr. *iwr<sup>ew</sup>* and a plural acc./loc. *iwr<sup>ans</sup>*, gen./dat./abl. *-eanc<sup>c</sup>*, instr. *-eambk<sup>c</sup>*. Traditionally the stem *iwr(-)* is derived from \**sewero-*, but this is formally problematic as \**-ewe-* would be expected to yield \**-e-* (cf. gen./dat./abl. *aler* ‘flour’ < \**-eweros* etc.), and it is difficult to provide a functional justification for a *vṛddhi* derivative \**sēwero-*. According to an alternative explanation by de Lamberterie (2014), *iwr* rather derives from \**setro-* (reflexive stem \**se-* beside \**swe-*), which, in the corresponding paradigm of the possessive pronoun, would be continued as *ewr*, gen. *iwr<sup>oy</sup>* with the same change of *ew* > *iw* in unaccented syllables as the loanword *ewl* (later *ivl*), *iwloy* ‘oil’, and *hewr* (later *hiwr*), *hiwroy* ‘stranger, guest’, allegedly a lexicalized variant of *iwr*. Another reflexive stem is *ink<sup>n</sup>* ‘-self’.

The function of reciprocal pronouns is fulfilled by the apparently reduplicated stems *mimeans* and *irears* (acc.) ‘each other’, the former presumably from \**smi-smia(h<sub>2</sub>)-*, cf. Lat. *alter . . . alterum* ‘one . . . the other’.

As in other Indo-European languages the possessive pronouns are *o*-stem adjectives: *im* ‘my’, *k<sup>c</sup>o* ‘your’, *mer* ‘our’, *jer* ‘your’. The forms are identical with the genitive of the personal pronouns.

### *The interrogative, indefinite and relative pronouns*

In the protolanguage the interrogative and indefinite pronoun were derived from an alternating stem \**k<sup>w</sup>e/o-* : \**k<sup>w</sup>i-*, cf. Ved. \**k<sup>w</sup>ás* > *káh* ‘who?’ vs. the indefinite particle *-cit* < \**k<sup>w</sup>id*; Lat. neuter adjectival *quod* vs. substantival *quid*, both interrogative and indefinite; or Av. gen. sg. *kahiiā* < \**k<sup>w</sup>osyo* beside *cahiiā* < \**k<sup>w</sup>esyō*. The original distribution between \**k<sup>w</sup>e/o-* and \**k<sup>w</sup>i-* is not entirely clear, but at least it seems likely that \**k<sup>w</sup>e-* is an archaism as opposed to \**k<sup>w</sup>o-* with the generalized *o*-vocalism of the thematic vowel known from the noun inflection. Despite the general loss of gender distinctions in Armenian the interrogative and indefinite pronouns distinguish between an animate ‘who?’, nom./acc. *o(v)*, gen. *oyr*, dat./loc. *um*, abl. *umē*, instr. *orov* and a neuter ‘what?’, nom./acc. *zi*, gen. *ēr*, dat./loc. (*h*)*im*, abl. *imē*, instr. *iw*. If \**k<sup>w</sup>* merges with \**p* before \**o* (cf. also *oīn* ‘spine, backbone’ < \**k<sup>w</sup>olso-*, Lat. *collus*, *collum*, Goth. *hals*; apparently no counterexamples), we are dealing with the stem \**k<sup>w</sup>o-* in the animate forms with gen. *oyr* < \**k<sup>w</sup>osyo-* + *-r* (Ved. *kásya*, Av. *kahiiā*) and dat./loc. *um* < \**k<sup>w</sup>o-sm-* (Ved. dat. *kásmai*, loc. *kásmin*). The neuter *z-i*, on the other hand, cannot go back to \**-k<sup>w</sup>id* (> \**(-)*č<sup>c</sup>(i), cf. the equation *oč<sup>c</sup>* ‘not’ = Gr. οὐκί < \**h<sub>2</sub>oyu-k<sup>w</sup>id* for \**ne h<sub>2</sub>oyu-k<sup>w</sup>id* ‘not in a lifetime’, cf. Cowgill 1960), so here we have to assume analogical absence of palatalization in analogy with the animate *ov*; the gen. *ēr* seems to continue \**k<sup>w</sup>esyō-* + *-r*, again with missing palatalization, but otherwise corresponding to Av. *cahiiā*. An adjectival variant *or* may go

back to \*k<sup>w</sup>otero- = Ved. *katará-*, Gr. *πότερος* (Schmitt 1981: 123). The same forms are also used in the function of relative pronouns, as, e.g., in the Italic languages.

The indefinite *ok<sup>c</sup>* ‘someone’ and *-k<sup>c</sup>* in the phrase *č<sup>c</sup>-ik<sup>c</sup>* ‘there is not’ are constructed by the addition of *-k<sup>c</sup>* < \**-k<sup>w</sup>e* to the stems *o-* and *i-* of the interrogative pronoun, cf. Ved. *kás-ca* ‘someone’, Lat. *quisque* ‘everyone’. Another indefinite pronoun is *omn* ‘someone’.

### *Demonstrative pronouns*

A striking feature of Armenian morphology is the consistent distinction between three demonstrative markers with first, second and third person deixis, *s* ‘hic’, *d* ‘iste’ and *n* ‘ille’ from \**ko-*, \**to-* and \**no-* respectively. These stems are the basis of the enclitic postponed articles *-s*, *-d*, *-n*; the anaphoric *sa*, *da*, *na* (probably from \**so-ay*, \**do-ay*, \**no-ay*); the demonstrative *ays*, *ayd*, *ayn*; the pronouns expressing identity *soyn*, *doyn*, *noyn* ‘this, that same’; various adverbs such as *aysr*, *aydr*, *andr* ‘here, hither; there, thither’, *asti*, *ayti*, *anti* ‘from here, from there’, *ayspēs*, *aydpēs*, *aynpēs* ‘in this, that way’; the corresponding adjectives *ayspisi*, *aydpisi*, *aynpisi* ‘of this, that kind’; and the interjections *a(ha)wasik*, *a(ha)wadik*, *a(ha)wanik* ‘look here/there’.

The anaphoric pronouns, here exemplified by *sa* (similarly *da* and *na*), are inflected as follows:

### The anaphoric pronoun *sa*

Singular	nom./acc.	<i>s-a</i>	Plural	nom.	<i>sok<sup>c</sup>-a</i>
	gen.	<i>sor-a</i>		acc./loc.	<i>sos-a</i>
	dat./loc.	<i>sm-a</i>		gen./dat./abl.	<i>soc<sup>c</sup>-a</i>
	abl.	<i>sm-anē</i>		instr.	<i>sok<sup>c</sup>-awk<sup>c</sup></i>
	instr.	<i>sov-aw</i>			

The basic stems are *so-*, *do-*, *no-* with “internal inflection” where the genitives *sora*, *dora*, *nora* may be analyzed as \**-otero-* + the enclitic particle *-ay*, and the dat./loc. and abl. sg. as \**so-sm-*.

### Numerals

The cardinal numbers from one to ten can all be derived from their Indo-European proto-forms with a few minor adjustments: *mi* ‘one’ < \**smih<sub>2</sub>*, the old feminine, cf. Gr. *μία*; *erku* ‘two’ < \**dwō*, Lat. *duo* etc.; *erek<sup>c</sup>* ‘three’ < \**treys*, Lat. *trēs*, Ved. *tráyah*; *č<sup>c</sup>ork<sup>c</sup>* ‘four’ < \**k<sup>w</sup>etores*, dissimilated from \**k<sup>w</sup>etwores*, Ved. *catvārah*; *hing* ‘five’ < \**penk<sup>w</sup>e*, Gr. *πέντε*, Ved. *pāñca*; *vec<sup>c</sup>* ‘six’, perhaps from a Lindeman variant \**suweks* (Klingenschmitt 1982: 61), cf. W *chwech*; *ewt<sup>c</sup>n* (with the variant *eawt<sup>c</sup>n*) ‘seven’ < \**septm<sub>7</sub>*, Lat. *septem* etc.; *ut<sup>c</sup>* ‘eight’, probably from \**optō* rather than \**oktō* under the influence of ‘seven’; the somewhat enigmatic *inn* ‘nine’, Gr. *ἐννέα* vs. Lat. *nouem* etc., whether from a metathesized \**enun* as assumed by Eichner (1978: 152) or actually exhibiting a laryngeal-based prothetic vowel; and, finally, *tasn* ‘ten’ < \**dek<sub>10</sub>* with unclear vocalism. The singular *mi*, which may also have the function of an indefinite article, has an aberrant inflection with gen. *mioj* beside *mioy* (cf. *kin*, *knoj* ‘woman’, also a stem originally ending in \**-h<sub>2</sub>*). Apart from *k<sup>c</sup>san* ‘20’, Lat. *uīginti* etc., the decadic numerals are characterized by the

element *-sun* < \*-(d)kōmt<sub>2</sub>, Gr. -kovta, in some cases with compensatory lengthening of the preceding element: *eresun* (< \*eria-) ‘30’; *k<sup>c</sup>arasun* ‘40’ as if < \*k<sup>w</sup>tw<sub>1</sub>, cf. Lat. *quadrāginta*; the formally remarkable *yisun* ‘50’ beside *hing* ‘five’; *vat<sup>c</sup>sun* ‘60’ beside *vec<sup>c</sup>* ‘six’; *ew<sup>c</sup>anasun* ‘70’; *ul<sup>c</sup>sun* ‘80’; and *innsun* ‘90’. However, the inherited word for ‘100’, \*k<sup>w</sup>mtóm, has been replaced by the etymologically unclear *hariwr*, and the Iranian loanwords *hazar* and *biwr* have been introduced for the larger numbers ‘1000’ and ‘10,000’.

Ordinals, apart from *arajin* ‘first’ < \*p<sub>1</sub>h<sub>3</sub>wyo- + *-ino-* (cf. Ved. adv. *pūrvyám* ‘first’) are derived with the suffixes *-(i)r*, *-ord*, or the combination *(-e)rord*: *erkir* and *erkrord* ‘second’, *erir* and *errord* ‘third’, *č<sup>c</sup>orir* and *č<sup>c</sup>orrord* ‘fourth’, *hngerord* ‘fifth’, with preservation of the final *-e-* of \*penk<sup>w</sup>e and hence *vec<sup>c</sup>erord* ‘sixth’. *Erkir* and *erir* also have the function of numeral adverbs ‘in the second place, zweitens’, ‘in the third place, drittens’, perhaps from the adverbial \*dwis, \*tris ‘twice, three times’ with RUKI development of \*-s rather than the otherwise isolated protoforms \*dwiro-, \*triro-, cf. Ved. *dviḥ*, *triḥ*; Gr. δίς, τρίς, Lat. *bis*, *ter*. The semantic development would be ‘two, three times’ → ‘the second, third time’ → ‘second, third’, cf. also the partly competing *erkic<sup>c</sup>s* ‘twice, the second time’ and *eric<sup>c</sup>s* ‘three times, the third time’. As for the productive suffix *-ord*, the most reasonable reconstruction may be \*k<sup>w</sup>ortos, cf. Lith. *kařtas*, OCS *kratъ* ‘time’ (Winter 1992) or even \*-k<sup>w</sup>rt as in Ved. *sak<sup>1</sup>t*, Av. *hak<sup>1</sup>arəḷ* ‘once’ (with special development of the sonant after labial/labiovelar), in particular in consideration of the alternative function of *-ord* as an agent noun suffix, e.g. *ors-ord*, ‘hunting-maker’, i.e. ‘hunter, fisher’: Ved. *mantra-k<sup>1</sup>t* ‘mantra-maker’ (Olsen 1999: 527–529).

## Verbs

The Armenian verb has two aspect stems (imperfective and aorist), two verbal voices (active and mediopassive), two verbal tenses (present and preterite), three modes (indicative, optative and imperative), three persons (1, 2 and 3) and two numbers (singular and plural). Thus, the perfect, the distinction between optative and subjunctive, the injunctive as a specific category, and the dual have been lost.

With a distinction between four vocalic stem types, *e-*, *i-*, *a-* and *u-*stems (the single *o-*stem *gom* ‘be, exist’, is based on an old perfect \*-h<sub>2</sub>wos-), the synchronic formation of the present indicative is entirely regular, even including the verb *em* ‘be’, which is identical with the endings of an *e-*verb:

### The present indicative

#### The present indicative

	<i>e-</i> stems		<i>i-</i> stems	<i>a-</i> stems	<i>u-</i> stems
	<i>em</i> ‘I am’	<i>berem</i> ‘I carry’	<i>berim</i> ‘I am carried’	<i>lam</i> ‘I weep’	<i>helum</i> ‘I pour’
sg. 1	<i>em</i>	<i>berem</i>	<i>berim</i>	<i>lam</i>	<i>helum</i>
2	<i>es</i>	<i>beres</i>	<i>beris</i>	<i>las</i>	<i>helus</i>
3	<i>ē</i>	<i>berē</i>	<i>beri</i>	<i>lay</i>	<i>helu</i>
pl. 1	<i>emk<sup>c</sup></i>	<i>beremk<sup>c</sup></i>	<i>berimk<sup>c</sup></i>	<i>lamk<sup>c</sup></i>	<i>helumk<sup>c</sup></i>
2	<i>ēk<sup>c</sup></i>	<i>berēk<sup>c</sup></i>	<i>berik<sup>c</sup></i>	<i>layk<sup>c</sup></i>	<i>heluk<sup>c</sup></i>
3	<i>en</i>	<i>beren</i>	<i>berin</i>	<i>lan</i>	<i>helun</i>

The paradigm seems to have arisen through a compromise between the athematic verb *em* and the thematic inflection with generalization of the thematic vowel \*-e-, thus:

- 1 sg. \*h<sub>1</sub>es-mi → *em* (for \*im) ⇒ *berem*  
 2 sg. \*h<sub>1</sub>es-si > *es* ⇒ *beres* (for \*berē)  
 3 sg. \*b<sup>h</sup>ere-ti > *berē* ⇒ *ē* (for \*est?)  
 1 pl. \*b<sup>h</sup>eromes → *beremk<sup>c</sup>* (analogical stem vowel) ⇒ *emk<sup>c</sup>*  
 2 pl. \*b<sup>h</sup>eretes > *berēk<sup>c</sup>* ⇒ *ēk<sup>c</sup>*  
 3 pl. \*h<sub>1</sub>senti (> \*in) → *en*; \*b<sup>h</sup>eronti → *beren* (analogical stem vowel)

Similarly the *a*- and *u*-stems. In the mediopassive *i*-inflection the stem vowel goes back to the stative marker \*-eh<sub>1</sub>-.

### *The imperfect*

In the imperfect there is no distinction between active and mediopassive, thus *berēi* ‘I was carrying’, ‘I was being carried’:

### **The imperfect**

	Singular	Plural
1	<i>berēi</i>	<i>berēak<sup>c</sup></i>
2	<i>berēir</i>	<i>berēik<sup>c</sup></i>
3	<i>berēr</i>	<i>berēin</i>

Similarly, *ei* ‘I was’ etc., from *a*-verbs *layi* ‘I was weeping’, from *u*-verbs *helui* ‘I was pouring’ etc. Here the endings are much less transparent than in the present, and their historical background is a matter of debate. The *-i*- endings of the singular are variously derived from either the imperfect \*e-h<sub>1</sub>es- (Jasanoff 1979; Klingenschmitt 1982) or the athematic optative (Winter 1975) with reference to English constructions like ‘he would carry’. In particular, the 2 sg. ending (-)ir may be a direct continuation of \*e-h<sub>1</sub>es-s under the assumption of a regular RUKI development (Olsen 1989a: 9).

### *The imperfective (“present”) imperative*

The imperative based on the imperfective stem is only used in the 2 sg. and pl. with the particle *mi* (cf. Ved. *mā́*, Gr. *μή*) in the function of a prohibitive. Formally, the 2 sg. is characterized by the ending *-r*, while the 2 pl. is identical with the indicative, e.g. 2 sg. *berer*, 2 pl. *berēk<sup>c</sup>*. If the enigmatic *r*-ending does not go back to the athematic \*-d<sup>hi</sup> as suggested by Jasanoff (1979: 146), it may have been transferred from aorist imperatives of the type *dir* ‘put’, *tur* ‘give’ from the old aorist injunctives \*d<sup>h</sup>eh<sub>1</sub>-s, \*deh<sub>3</sub>-s with RUKI development of the final \*-s (Olsen 1989a: 9).

### *The aorist indicative*

The Armenian aorist stems are found in two variants, the so-called root and weak aorist. While the root aorist is based on the synchronic root and may thus go back to either an old imperfect or an aorist stem, e.g. 3 sg. *e-ber* ‘(s)he carried’ < impf. pret. \*e-b<sup>h</sup>er-e-t, Ved.

*á-bhar-a-t* vs. *e-t* ‘(s)he gave’ < aor. \**e-deh<sub>3</sub>-t*, Ved. *á-dāt*, the weak aorist is characterized by a morpheme *-c<sup>c</sup>-*, in *e*-verbs *-ec<sup>c</sup>-*, 3 sg. *-eac<sup>c</sup>*. In the aorist the mediopassive is distinguished from the active by a separate set of verbal endings, e.g. *beri* ‘I carried’, *sirec<sup>c</sup>i* ‘I loved’ vs. *beray* ‘I was carried’, *sirec<sup>c</sup>ay* ‘I was loved’:

### The aorist indicative

#### Active

sg. 1	<i>beri</i>	<i>sirec<sup>c</sup>i</i>
2	<i>berer</i>	<i>sirec<sup>c</sup>er</i>
3	<i>e-ber</i>	<i>sireac<sup>c</sup></i>
pl. 1	<i>berak<sup>c</sup></i>	<i>sirec<sup>c</sup>ak<sup>c</sup></i>
2	<i>berēk<sup>c</sup>, -ik<sup>c</sup></i>	<i>sirec<sup>c</sup>ēk<sup>c</sup>, -ik<sup>c</sup></i>
3	<i>berin</i>	<i>sirec<sup>c</sup>in</i>

#### Mediopassive

sg. 1	<i>beray</i>	<i>sirec<sup>c</sup>ay</i>
2	<i>berar</i>	<i>sirec<sup>c</sup>ar</i>
3	<i>beraw</i>	<i>sirec<sup>c</sup>aw</i>
pl. 1	<i>berak<sup>c</sup></i>	<i>sirec<sup>c</sup>ak<sup>c</sup></i>
2	<i>berayk<sup>c</sup></i>	<i>sirec<sup>c</sup>ayk<sup>c</sup></i>
3	<i>beran</i>	<i>sirec<sup>c</sup>an</i>

Apart from the endingless 3 sg. going back to \*-*et* and supplied with the augment in what would otherwise be a monosyllabic form, the active endings are similar to those of the imperfect. The mediopassive is characterized by the vowel *-a-*, which would be regular in the 3 pl. \*-*nto* > *-an* and the 3 sg. of *seř* roots \*-*ato* > *-aw*, e.g. *cnaw* (< \**cinaw*) ‘begot, gave birth’, Gr. ἐγένετο.

While neither the consonantal element *-c<sup>c</sup>-* nor the segment *-ea-* (> *-e-*) of the weak aorist is fully understood, there exists a more archaic-looking type where *-c<sup>c</sup>-* directly follows the root vowel, thus *lc<sup>c</sup>i* ‘I filled’, 3 sg. *e-lic<sup>c</sup>* (root \**pleh<sub>1</sub>-*), *bac<sup>c</sup>i* ‘I opened’ (root \**b<sup>h</sup>eh<sub>2</sub>-*). Traditionally *-c<sup>c</sup>-* is derived from \*-*sk-*, but a connection with the *s*-aorist may be more likely, cf. *lc<sup>c</sup>i*: Gr. ἐπλησα or the type in *-ac<sup>c</sup>i* compared to Greek denominative *s*-aorists like ἐτίμησα ‘I honored’. According to Klingenschmitt (1982: 286ff.), we are dealing with a “strengthened” \*-*ss-* > *-c<sup>c</sup>-* despite the apparent counterevidence of \**h<sub>1</sub>essi* > *es* ‘you are’. At any rate it is remarkable that the source of the *-c<sup>c</sup>-*aorist seems to be roots and stems ending in a postvocalic laryngeal, so as a variant of Klingenschmitt’s theory one might perhaps suggest a development \*-*Hs-* > *-c<sup>c</sup>-*, similar to \*-*Ks-* > *-c<sup>c</sup>-* under the assumption of “laryngeal hardening”.

#### The aorist imperative

As opposed to the prohibitive imperfective/“present” imperative, the aorist imperative is only used in positive orders. When the aorist is of the “root” type, the corresponding imperative active 2 sg. is identical to the naked root (diachronic stem), e.g. *ber* ‘carry’ < \**b<sup>h</sup>ere* (aor. *beri*, 3 sg. *e-ber*), while the 2 pl. vacillates between the endings *-ēk<sup>c</sup>* (< \*-*etes*) and *-ik<sup>c</sup>*. The 2 sg. imperative of the weak aorist equals the aorist stem without the aorist marker *-c<sup>c</sup>(-)*, thus 3 sg. aor. ind. *sireac<sup>c</sup>*: 2 sg. aor. imp. *sirea* beside the 2 pl. *sirec<sup>c</sup>ēk<sup>c</sup>*, *sirec<sup>c</sup>ik<sup>c</sup>*. In



the mediopassive the forms of the root aorist are 2 sg. *ber/berir*, 2 pl. *beraruk<sup>c</sup>*, and those of the weak aorist are 2 sg. *sireac<sup>c</sup>*, 2 pl. *sirec<sup>c</sup>aruk<sup>c</sup>*, where *(-a-)ruk<sup>c</sup>* possibly derives from *\*(-a)d<sup>h</sup>uwe* + plural marker *-k<sup>c</sup>* (cf. Skr. mid. imp. *bharadhvam*), as suggested by Jasanoff (1979: 144f.), if *-r-* is a (conditioned) continuation of *\*-d<sup>h</sup>-*, possibly before *-u-*.

### *The subjunctives*

Armenian has two subjunctives, one derived from the imperfective (“present”) stem, the other from the aorist stem with the primary function of future. In both cases the subjunctive marker is *-ic<sup>c</sup>-* (unaccented *-c<sup>c</sup>-*).

The primary type is the aorist subjunctive with the following inflection:

### **The aorist subjunctive**

	active	mediopassive
sg. 1	<i>beric<sup>c</sup></i>	<i>berayc<sup>c</sup></i>
2	<i>berc<sup>c</sup>es</i>	<i>berc<sup>c</sup>is</i>
3	<i>berc<sup>c</sup>ē</i>	<i>berc<sup>c</sup>i</i> (< <i>*-iy</i> )
pl. 1	<i>berc<sup>c</sup>uk<sup>c</sup></i>	<i>berc<sup>c</sup>uk<sup>c</sup></i>
2	<i>berjik<sup>c</sup></i>	<i>berjik<sup>c</sup></i>
3	<i>berc<sup>c</sup>en</i>	<i>berc<sup>c</sup>in</i>

When the subjunctive marker is added to a weak aorist stem, *\*-c<sup>c</sup>-c<sup>c</sup>-* is dissimilated to *-s-c<sup>c</sup>-* and *\*-c<sup>c</sup>-j-* to *-s-j-*, e.g. from *sirem* ‘love’: *sirec<sup>c</sup>ic<sup>c</sup>*, *siresc<sup>c</sup>es*, *siresc<sup>c</sup>ē*, *siresc<sup>c</sup>uk<sup>c</sup>*, *siresjik<sup>c</sup>*, *siresc<sup>c</sup>en*. The ending *-uk<sup>c</sup>* of the 1 pl. may derive from *\*-omes*, the only case apart from the participial *-un* (< *\*-ont-/omə<sub>1</sub>no-*) where the *o*-alternant of the thematic vowel is preserved, while *-jik<sup>c</sup>* probably represents an old optative *\*-yeh<sub>1</sub>-tes* as suggested by Klingenschmitt (1982: 40).

In the imperfective (“present”) subjunctive, *-ic<sup>c</sup>-* is simply added to the stem vowel and followed by the present endings, thus *berem*: *\*bere-ic<sup>c</sup>-em* > *\*berēc<sup>c</sup>-em* > *beric<sup>c</sup>em* etc.; *berim*: *beric<sup>c</sup>im*; *lam*: *layc<sup>c</sup>em*; and *helum*: *heluc<sup>c</sup>um* with secondary adaptation to the *u*-stem inflection.

Traditionally the element *-c<sup>c</sup>-* is derived from the Indo-European suffix *\*-ske/o-* (e.g. Schmitt 1981: 143) with various explanations of the preceding *-i-*, whether as a match of Gr. *-ίσκω* in verbs like *εὑρίσκω* ‘find’ or a combination of the optative morpheme + *\*-sk-*. Alternatively, Lühr (1994) assumes a reinterpretation of the indicative *hayc<sup>c</sup>em* ‘ask (for)’ as a subjunctive. A different scenario is proposed by Olsen (1988): if the 3 sg. thematic subjunctive *\*-ē-ti*, the unmarked form of the paradigm, was reinterpreted as a stem and supplied with a new set of personal endings we would end up with the attested paradigm of the aorist subjunctive: *\*b<sup>h</sup>erēty-ō* > *beric<sup>c</sup>*, *\*b<sup>h</sup>erēty-essi* > *berc<sup>c</sup>es*, *\*b<sup>h</sup>erēty-eti* > *berc<sup>c</sup>ē*, etc. The pattern would be similar to that of, e.g., Modern Pers. *hast* ‘he is’ ⇒ *hastam* ‘I am’ etc.

### *Verbal stem formation*

To some extent the vocalic stem classes reflect inherited types:

- e*-verbs with corresponding mediopassives in *-i-* < stative *\*-eh<sub>1</sub>-*, e.g.:
- simple thematic verbs: *berem* ‘carry’, Ved. *bhārati*
- denominatives in *\*-eye-*: *srbem* ‘cleanse’ from *surb* ‘pure’

<i>y</i> -verbs :	<i>mrmnjem</i> ‘mumble’ ( <i>-nj-</i> < * <i>-ny-</i> )
* <i>-ske</i> -verbs:	<i>čanač'em</i> (assimilated from * <i>canač'em</i> ) ‘know’ < * <i>ğḡh<sub>1</sub>-ske-</i> , Gr. <i>γῆνωσκω</i>
inchoative statives:	<i>hangč'im</i> ‘rest’ < * <i>sḡ-k<sup>w</sup>i(e)h<sub>1</sub>-ske-</i> , Lat. <i>conquiēscō</i> (Klingenschmitt 1982: 78f.)
old perfects:	<i>gitem</i> ‘know’ < * <i>woyd-e-</i>
<i>a</i> -verbs, e.g.:	
root-verbs:	<i>bam</i> ‘speak’ < * <i>b<sup>h</sup>eh<sub>2</sub>-</i> , Lat. <i>fāri</i>
reduplicated verbs:	<i>tam</i> ‘give’ < * <i>di-də<sub>2</sub>-</i> , Gr. <i>δίδωμι</i> , 1 pl. <i>δίδομεν</i>

More complex formations include variations of nasal presents, e.g. *lk'anem* ‘leave’, Ved. *rinákti*, Lat. *linquō*, Gr. *λιμπάνω*; Ved. *sparnam* ‘threaten’, Lat. *spernō*; *z-genum* ‘dress’ < \**-wes-nu-*, Gr. *ἔννυμι*; and the productive type of denominatives in *-anam*, e.g. *ceranam* ‘grow old’. A particularly important subtype is the causatives in *-uc<sup>e</sup>anem*, aor. *-uc<sup>e</sup>i* (3 sg. *-oyc<sup>e</sup>*), e.g. *usanim* (aor. *usay*) ‘learn’, cf. Lith. *jũnkti* ‘get used to’, vs. causative *usuc<sup>e</sup>anem* ‘teach’. In the paradigmatic relation between imperfective and aorist stems we find traces of the inherited state of affairs as when a nasal present is matched by a “root” aorist, e.g. *bekanem* ‘break’, Ved. *bhanákti* : aor. *beki*, 3 sg. *e-bek*.

### *Non-finite verb forms*

The Armenian system of infinitives and participles (monographically treated by Stempel 1983) consists of an *o*-stem past participle in *-eal*, mostly based on the aorist stem, e.g. *beréal* ‘having carried’, cf. OCS *neslŭ jesmŭ* ‘I have carried’, a voice-indifferent *o*-stem infinitive in *-l* < \**-Vlo-*, e.g. *berel* ‘carry/be carried’ (from *berem/berim*), *lal* ‘weep’, *hehul* ‘pour’, which is the basis of an indeclinable future participle in *-eloc<sup>e</sup>*, e.g. *bereloc<sup>e</sup>* ‘to be carried’ and an *ea*-stem gerund in *-eli*, e.g. *sireli* ‘lovable, beloved’, formally and functionally matching Tocharian constructions like Toch. B *mā yokalle* ‘there is not to be drunk, one should not drink’. To these one may add the mostly substantivized verbal adjectives in *-un*, e.g. *sirun* ‘beloved’ and ‘loving’, probably a merger of the active \**-ont-* and the mediopassive \**-omə<sub>1</sub>no-*, and the agent nouns/active participles in *-ol/-awl*, e.g. *cnawl* ‘parent’ (see under “Nominal Word Formation”). As the suffix of the old passive participle in \**-to-* lost its phonetic transparency and was thus unsuitable as a marker of a whole morphological category, it left only isolated traces in lexicalizations such as *mard* (*o*-st.) ‘man’ < \**mṛto-* (isolated from the privative \**ḡ-mṛto-* ‘immortal’), *li* ‘full’ < \**pleh<sub>1</sub>-to-* and *lu* ‘heard (of)’, i.e. ‘well-known’ < \**klu-to-*.

## NOMINAL WORD FORMATION

Due to the dramatic phonetic changes, including loss of the vowels in final syllables, many inherited noun suffixes are rendered opaque. Thus, e.g., the Indo-European suffix \**-ti-*, used for the formation of verbal abstracts, is reflected as *-<sup>e</sup>(i)-* after *u*-diphthongs, e.g. *alawt<sup>e</sup>-k<sup>e</sup>* (*i*-st.) ‘prayer’ < \**sḡh<sub>2</sub>-ti-*, and as *-d(i)-* after sonants, e.g. *bard* (*i*-st.) ‘pile’, Ved. *bhṛtī-* ‘support’, Av. *bərəti-*, and is lost intervocally in, e.g., *lir* (*i*-st.) ‘fullness’ < \**pleh<sub>1</sub>-ti-s*, Ved. *prātī-*. As a countermeasure against this development, several suffixes are occasionally linked together to create more characteristic and easily recognizable suffixal chains, e.g. the extremely productive abstract suffix *-ut<sup>e</sup>iwn* < \**-e(h<sub>1</sub>)u-ti-h<sub>3</sub>on(h<sub>2</sub>)-* (cf. Lat. *-tiō*, *-tiōnis*), which may be derived from verbal, adjectival and nominal stems

alike, e.g. *spananem* ‘kill’ → *spanut<sup>c</sup>iwn* ‘killing’, *cer* ‘old’ → *cerut<sup>c</sup>iwn* ‘oldness’, *elbayr* ‘brother’ → *elbayrut<sup>c</sup>iwn* ‘brotherhood’.

Examples of inherited suffixes and suffix clusters based on inherited elements are:

- ac<sup>c</sup>i*: adjectives denoting descent or appurtenance, e.g. *galileac<sup>c</sup>i* ‘Galilean’, *k<sup>c</sup>alak<sup>c</sup>i* ‘citizen’ < \**-skyo/ah<sub>2</sub>-*, cf. Goth. *judawisks* ‘Jewish’
- awl* (and *-ol*): agent nouns/active participles, e.g. *cnawl* ‘parent’, probably < \**ǵenə<sub>1</sub>-tlo-*, cf. Gr. *γενέτωρ*, Lat. *genitor*; the connecting vowel *-a-* may originate in *set* roots like \**ǵenə<sub>1</sub>-*, or in derivatives based on denominative *a*-verbs, cf. Lat. *imperāri* ⇒ *imperātor*, or simply derive from Iranian bases such as *nmanawl* ‘similar’ from \**nimāna-*
- ean*: adjective suffix, typically used in patronymics of the type *Abrahamean*; according to Olsen (1992 and 1999: 385–391) from IE \**-i-h<sub>3</sub>no-* = Lat. *-īnus* etc.
- i*: adjective and noun suffix, e.g. *asui* ‘woollen’ from *asu* ‘wool’, *kogi* ‘butter’ = Ved. *gāvya-*, Av. *gaoiia-*, Gr. *-βοιός* ‘cow-’, Toch. B *kewiye* ‘butter’; IE \**-(i)yo/ah<sub>2</sub>-*
- ič<sup>c</sup>-*: agent nouns, e.g. *ergič<sup>c</sup>* ‘singer’; IE \**-ikyah<sub>2</sub>-*, cf. Slavic derivatives in *-ač<sup>c</sup>* and *-ič<sup>c</sup>* such as OCS *kovač<sup>c</sup>* ‘blacksmith’, *kotorič<sup>c</sup>* ‘quarrelsome’
- oyt<sup>c</sup>*: abstract nouns derived from verbal or nominal stems, e.g. *erewoyt<sup>c</sup>* ‘appearance’ from *erewim* ‘appear’ (Gr. *πρέπω*); identical with the Greek suffix *-ευσίς* and the basis of the suffix conglomerate *-ut<sup>c</sup>iwn*
- umn*: action nouns, e.g. *gelumn* ‘rolling’, cf. Gr. *εἰλῦμα*, Lat. *uolūmen*
- un*: verbal adjectives, originally participial, but mostly substantivized, e.g. *zerun* ‘creeping’, i.e. ‘snake’; *anasun* ‘not talking’, i.e. ‘animal’. Since Meillet (1928: 2), traditionally interpreted as the middle participle corresponding to Gr. *-μενός*, but, as pointed out by Stempel (1983: 81–83), the function is often active, and the active participle in \**-ont-* would give the same result. Most likely, we are dealing with a merger of the two forms.

Numerous other common suffixes or original final members of compounds are borrowed from Iranian (cf. Greppin 1973, Jahowkyan 1993 and Olsen 1999). Examples of Iranian loan suffixes and suffix clusters are:

- agoyn*: adjectives denoting manner and appearance; intensive and comparative suffix, from MIran. *-gaona-*, e.g. *vardagoyn* ‘rosy’, Sogd. *wrdγwn*
- ak*: diminutives, e.g. *naw-ak* ‘small ship’
- akan*: adjective suffix, e.g. *gišer-akan* ‘nightly’
- apet*: nouns with the meaning ‘master of . . .’, from Iran. \**-pati-*, e.g. *k<sup>c</sup>alak<sup>c</sup>apet* ‘ruler of a city’
- astan*: local nouns, e.g. *herāstan* ‘a place far away’ from *heri* ‘distant’
- aran*: nouns denoting holders or containers, from Iran. \**-dāna/i-*, e.g. *ganjaran* ‘treasury’
- ik*: adjectives of appurtenance and diminutives, e.g. *hayrik* ‘little, dear father’
- pēs*: adverbs of manner, e.g. *ayspēs* ‘in this way’, cf. Av. *-paēsa(h)-*

### Nominal compounds

Most categories of compounds as known from other Indo-European languages such as Indo-Iranian, Greek and Germanic are maintained as living, productive types in Armenian,

where only prepositional compounds and compound verbs are rarely found. In many cases we are dealing with simple loans from an Iranian source or calques/semi-calques from either Iranian or Greek, e.g. Gr. μητρόπολις → *mayrak'alak'* 'mother city', but there is ample evidence for indigenous creations such as *tn-ank* 'poor' ('(whose) house (*tun*) (has) fallen'). The following examples will serve to illustrate the most common subtypes:

*Dependent determinatives*: *tēr* 'master' (*ti-* < \**dems-* + *ayr* 'man', cf. Gr. δεσπότης, Ved. *dám-pati-*; *tikin* 'lady, mistress' (*ti-* + *kin* 'woman'); *skesrayr* 'father-in-law' (*skesur* 'mother-in-law' + *ayr* 'man, husband')

*Possessive compounds* (*bahuvrīhis*): *jēmunayn* 'empty-handed' (lit. 'hand-empty'); *č'arakn* 'envious', lit. 'evil-eye', a calque from Miran., cf. MPers. *duščašm*

*Verbal governing compounds*, active or passive meaning: *dimadarj* 'opponent', lit. 'face-turning'; *amawt'alic'* 'shameful'.

## SYNTAX

In general, Classical Armenian syntax is similar to that of other ancient Indo-European languages with preservation of a rich case system. The word order is relatively free, with SVO as the unmarked pattern. In translations from Greek originals it has to a large extent been possible to maintain the Greek word order. A few characteristic points deserve to be mentioned:

- the concordance between adjectives and substantives in case and number is not always observed. Most often a preceding adjective is left uninflected, e.g. *bazum awurk'* but *awurk' bazumk'* 'many days' (Meillet 1936: 137).
- the definite object is marked by a preceding *z-*, reminiscent of the Middle Parthian use of *ō* < \**h<sub>2</sub>m̥b<sup>h</sup>i* in connection with both the indirect and the direct object.
- a peculiar construction, the "transitive perfect" is found in connection with the participle in *-eal*, where the agent is in the genitive and the object in the accusative, including the marker *z-*, e.g. *nora gorceal ē z-gorc* 'he has done the work'.

## FURTHER READING

The literature on Armenian historical grammar in English is rather scarce, but for a general introduction one may refer to Godel's sketch (1970) and his concise monograph (1975). A good historical introduction to the classical language in German, including text samples, is Schmitt 1981, and for a contemporary French introduction de Lamberterie 1988–89 is recommended. Otherwise, the brief, but wonderfully clear classic grammar by Meillet (1936) remains an unsurpassed treasure trove of information.

A comprehensive etymological dictionary is still a desideratum. Evidently Hübschmann's pioneering work (1897) is severely outdated, and Ačaryan's four-volume dictionary in Armenian is practically inaccessible to the beginner. To some extent the need has been satisfied by Martirosyan's recent and in many ways excellent work (2010), but unfortunately only part of the inherited vocabulary is covered. As a supplement one may consult the lexical study by Solta (1960).

Monographic studies on specific subjects include Klingenschmitt 1982 on the verb, Olsen 1999 on the noun and Matzinger 2005 on nominal inflection. Holst 2009 is a highly idiosyncratic treatment of selected phonological and morphological problems.

The important early articles on Armenian by Pedersen (1982) and Meillet (1977) have now been made more easily accessible in collected volumes, and the same goes for Kortlandt's very original work on the subject (2003).

## REFERENCES

- Ačařyan, Hr. 1926–35 (1971–79), *Hayeren Armatakan Bararan* I-IV, Erevan
- Beekes, Robert S.P. 1988, "Laryngeal developments: A survey," in: Bammesberger, Alfred (ed.), *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*, Carl Winter, Heidelberg: 59–105
- Bolognesi, Giancarlo 1960, *Le fonti dialettali degli imprestiti iranici in armeno*, Milano, *Società edivice(?) vita e peusiero*
- Clackson, James 1994, *The Linguistic Relationship between Armenian and Greek*, Publications of the Philological Society 30, Oxford–Cambridge
- Cowgill, Warren 1960, "Greek *ou* and Armenian *oč*," *Language* 36: 347–350
- de Lamberterie, Charles 1988–89, "Introduction à l'arménien classique," *LALIES* 10, 233–289
- de Lamberterie, Charles 2014, "Armeniën *iwr* et *hiwr*, grec *ἐταῖρος*: le 'soi' et le 'sien', l' 'hôte' et le 'compagnon'," in: Martirosian, Aram; Ouzouhian, Angès & Zuckerman, Constantin (eds.), *Mélanges Jean-Pierre Mahé, Travaux et Mémoires* 18, Paris: 429–438
- Dumézil, Georges 1938, "Traitement de *m + p* en arménien," *Bulletin de la Société de Linguistique de Paris* 39: 241f
- Džaukjan, G. B. 1982, *Sravitel' naja grammatika armjanskogo jazyka*, Erevan, Academy Press
- Eichner, Heiner 1978, "Die urindogermanische Wurzel \**Hreu* 'hell machen'," *Sprache* 24/2: 144–162
- Francis, Eric D. 1970, *Greek disyllabic roots: The aorist formations* [PhD Dissertation, Yale]
- Gamkrelidze, Thomas 1985, "The Indo-European 'glottalic theory' and the system of Old Armenian consonantism," in: Pieper, U. & Stickel, G. (eds.), *Studia linguistica diachronica et synchronica Werner Winter sexagenario anno MCMLXXX*, Mouton de Gruyter, Berlin–New York–Amsterdam: 281–284
- Godel, Robert 1970, "Diachronic Armenian," in: Thomas Sebeok (ed), *Current Trends in Linguistics*, Mouton, The Hague–Paris: 139–159
- Godel, Robert 1975, *An Introduction to the Study of Classical Armenian*, Reichert, Wiesbaden
- Greppin, John A.C. 1973, "The Middle Persian suffixes in Classical Armenian," *Revue des Études Arméniennes* 10: 1–10
- Greppin, John A.C. 1993, "The development of \**Ty* in classical armenian," in: *The Second International Symposium of Armenian Linguistics (21–23 September 1987)*, Proceedings, 2, Erevan: 15–24
- Hamp, Eric P. 1988, "On the essentiality of the Armenian Nasal Declensions," *Annual of Armenian Linguistics* 9: 19f
- Holst, Jan Henrik 2009, *Armenische Studien*, Harrassowitz, Wiesbaden
- Hübschmann, Heinrich 1972<sup>3</sup>, *Armenische Grammatik, Erster Teil, Armenische Etymologie*, Georg Olms Verlag, Hildesheim–New York [first print Leipzig 1897]
- Ĵahowkjan, G.B. 1993, "The Armenian suffixes of Iranian origin," in: Brogyanyi, Bela & Lipp, Reiner (eds.), *Comparative-Historical Linguistics: Indo-European and Finno-Ugric. Papers in honor of Oswald Szemerényi*, John Benjamins, Amsterdam–Philadelphia: 257–269

- Jasanoff, Jay H. 1979, "Notes on the Armenian personal endings," *KZ* 93: 133–149
- Jensen, Hans 1959, *Altarmenische Grammatik*, Carl Winter, Heidelberg
- Karst, Josef 1901, *Historische Grammatik des Kilikisch-Armenischen*, Strassburg
- Klingenschmitt, Gert 1982, *Das altarmenische Verbum*, Reichert, Wiesbaden
- Klingenschmitt, Gert 1994, "Die Verwandtschaftsverhältnisse der indogermanischen Sprachen," in: Rasmussen, Jens Elmegård (ed.), *In honorem Holger Pedersen, Kolloquium der indogermanischen Gesellschaft vom 25. bis 28. März 1993 in Kopenhagen*, Reichert, Wiesbaden: 235–251
- Kloekhorst, Alwin 2008, *Etymological Dictionary of the Hittite Inherited Lexicon*, Brill, Leiden–Boston
- Kortlandt, Frederik 1975, "A note on the Armenian palatalization," *KZ* 89: 43–45
- Kortlandt, Frederik 1980a, "On the relative chronology of Armenian sound changes," in: Greppin, J.A.C. (ed.), *Proceedings of the First International Conference of Armenian Linguistics*, Caravan Books, Delmar (New York): 79–106
- Kortlandt, Frederik 1980b, "H<sub>2</sub>O and oH<sub>2</sub>," *Lingua Posnaniensis* XXIII: 127f
- Kortlandt, Frederik 1984, "Proto-Armenian case endings," in: Hovsepian, L., Parnasian, N. & Simonian, S. (eds.), *Proceedings of the Second International Symposium on Armenian Linguistics 2*, Erevan, Academy Press: 97–105
- Kortlandt, Frederik 1987, "Notes on Armenian historical phonology V," *Studia Caucasica* 7: 61–65
- Kortlandt, Frederik 1989, "The making of a puzzle," *Annual of Armenian Linguistics* 10: 43–52
- Kortlandt, Frederik 2003 (ed. R.S.P. Beekes), *Armeniaca – Comparative Notes*, Caravan Books, Delmar (New York)
- Lühr, Rosemarie 1994, "Zum Konjunktivmorphem *-ic<sup>c</sup>* im Armenischen," in: Rasmussen, Jens Elmegård (ed.), *In honorem Holger Pedersen, Kolloquium der Indogermanischen Gesellschaft vom 26. bis 28. März 1993 in Kopenhagen*, Reichert, Wiesbaden: 259–273
- Martirosyan, Hrach K. 2010, *Etymological Dictionary of the Armenian Inherited Lexicon*, Brill, Leiden–Boston
- Matzinger, Joachim 1992, *Der Armenische Reflex von idg. \*-u-*, Vienna
- Matzinger, Joachim 1998, Review of Vladimir Orël: A concise historical grammar of the Albanian language (Leiden – Boston – Köln 2000), *Sprache* 40: 109–123
- Matzinger, Joachim 2005, *Untersuchungen zum altarmenischen Nomen: Die Flexion des Substantivs*, J.H. Röhl, Dettelbach
- Meillet, Antoine 1903, "Observation sur la graphie de quelques anciens manuscrits de l'Évangile arménien," *Journal Asiatique* II: 487–507
- Meillet, Antoine 1924, "Remarques éymologiques," *Revue des Études Arméniennes* IV: 1–6
- Meillet, Antoine 1928, "Les noms du type *t<sup>c</sup>r̄çun*," *Revue des Études Arméniennes* IV: 1–3
- Meillet, Antoine 1936, *Esquisse d'une grammaire comparée de l'arménien classique*, Imprimerie des pp. mékhitaristes, Vienne
- Meillet, Antoine 1977, *Études de linguistique et de philologie arméniennes*, Imprimerie Orientaliste, Louvain
- Normier, Rudolf 1977, "Idg. Konsonantismus, germ. 'Lautverschiebung' und Verner'sches Gesetz," *KZ* 91: 171–218
- Normier, Rudolf 1980, "Tocharisch *ñkät/ñakte* 'Gott'," *KZ* 94: 251–281
- Olsen, Birgit Anette 1985, "On the development of Indo-European prothetic vowels in Classical Armenian," *Revue des Études Arméniennes* 19: 5–17

- Olsen, Birgit Anette 1986, "The Armenian continuations of Indo-European intervocalic \*w," *Annual of Armenian Linguistics* 7: 51–56
- Olsen, Birgit Anette 1988, "The formation of the subjunctive paradigm in Classical Armenian," *Annual of Armenian Linguistics* 9: 9–17
- Olsen, Birgit Anette 1989a, "Three notes on Armenian historical phonology," *Annual of Armenian Linguistics* 10: 5–25
- Olsen, Birgit Anette 1989b, "A trace of Indo-European accent in Armenian," *Historische Sprachforschung* 102: 220–240
- Olsen, Birgit Anette 1992, "The development of high vowel plus laryngeal in Armenian," in: Greppin, John A.C. (ed.), *Proceedings of the Fourth International Conference on Armenian Linguistics, Cleveland 1991*, Caravan Books, Delmar (New York): 129–146
- Olsen, Birgit Anette 1993, "On the origins of Armenian c<sup>e</sup>," in: *The Second International Symposium of Armenian Linguistics (21–23 September 1987)*, *Proceedings* 2, Erevan, Academy Press: 25–43
- Olsen, Birgit Anette 1999, *The Noun in Biblical Armenian*, Mouton–De Gruyter, Berlin–New York
- Olsen, Birgit Anette 2005, "The development of IE \*mp and \*mb<sup>h</sup> in Armenian – Dumézil reconsidered," in: Schweiger, Günter (ed.), *Indogermanica, Festschrift Gert Klingenschmitt. Indische, iranische und indogermanische Studien dem verehrten Jubilar dargebracht zu seinem fünfundsechzigsten Geburtstag*, Taimering, Schweiger VWT–Verlag: 443–443
- Olsen, Birgit Anette 2009, "The Conditioning of Laryngeal Beaking in Greek," in: Lühr, Rosemarie & Ziegler, Sabine (eds.), *Protolanguage and Prehistory, Akten der XII. Fachtagung der Indogermanischen Gesellschaft vom 11. bis 15. Oktober 2004 in Krakau*, Reichert, Wiesbaden: 348–365
- Olsen, Birgit Anette 2010, *Derivation and Composition – Two Studies in Indo-European Word Formation*, Innsbrucker Beiträge zur Sprachwissenschaft (Band 136), Innsbruck
- Pedersen, Holger 1905, "Zur armenischen sprachgeschichte," *KZ* 26: 277–340
- Pedersen, Holger 1982, *Kleine Schriften zum Armenischen* [Herausgegeben von Rüdiger Schmitt], Georg Olms, Hildesheim–Zürich–New York
- Pisani, Vittore 1950–51, "Sulla fonetica dell'armeno," *Ricerche Linguistiche* 1: 165–193, 2: 47–74
- Schmitt, Rüdiger 1981, *Grammatik des Klassisch-Armenischen*, Innsbrucker Beiträge zur Sprachwissenschaft, Innsbruck
- Schmitt, Rüdiger 1983, "Iranisches Lehngut im Armenischen," *Revue des Études Arméniennes* 17: 73–112
- Seldeslachts, Herman 1991, "Varia Etymologica," in: Isebaert, Lambert (ed.), *Studia Etymologica Indoeuropaea Memoriae A.J. van Windekens Dicata*, Orientalia Lovaniensia Analecta, Leuven: 247–261
- Solta, Georg Rhenatus 1960, *Die Stellung des Armenischen im Kreise der indogermanischen Sprachen. Eine Untersuchung der indogermanischen Bestandteile des armenischen Wortschatzes*, Mechitaristenverlag, Wien
- Solta, Georg Rhenatus 1990, "Die Stellung des Armenischen im Kreise der indogermanischen Sprachen. Ein Überblick," in: Ruprechtsberger, Erwin M. (ed.), *Armenien, Beiträge zur Sprache, Geschichte und Archäologie*, Stadtmuseum Linz: 7–17
- Stempel, Reinhard 1983, *Die Infinite Verbalformen des Armenischen*, Peter Lang, Frankfurt am Main–Bern–New York
- Vennemann, Theo 1986, "Syllable-based sound changes in early Armenian," *Annual of Armenian Linguistics* 7: 27–43

- 
- Viredaz, Rémy 2003, "Notes on Armenian historical phonology," *Annual of Armenian Linguistics* 24–25: 85–103
- Weitenberg, J.J.S. 1989, "The inflections of proper names in Armenian," *Annual of Armenian Linguistics* 10: 57–72
- Winter, Werner 1965, "Evidence in Armenian," in: Winter, Werner (ed.), *Evidence for Laryngeals*, Mouton, London–The Hague–Paris
- Winter, Werner 1975, "Die Personalendungen des Imperfekts und Aorists im Armenischen," *KZ* 89: 110–122
- Winter, Werner 1992, "Armenian," in: Gvozdanović, J. (ed.), *Indo-European Numerals*, Mouton de Gruyter, Berlin–New York: 347–359



# TOCHARIAN

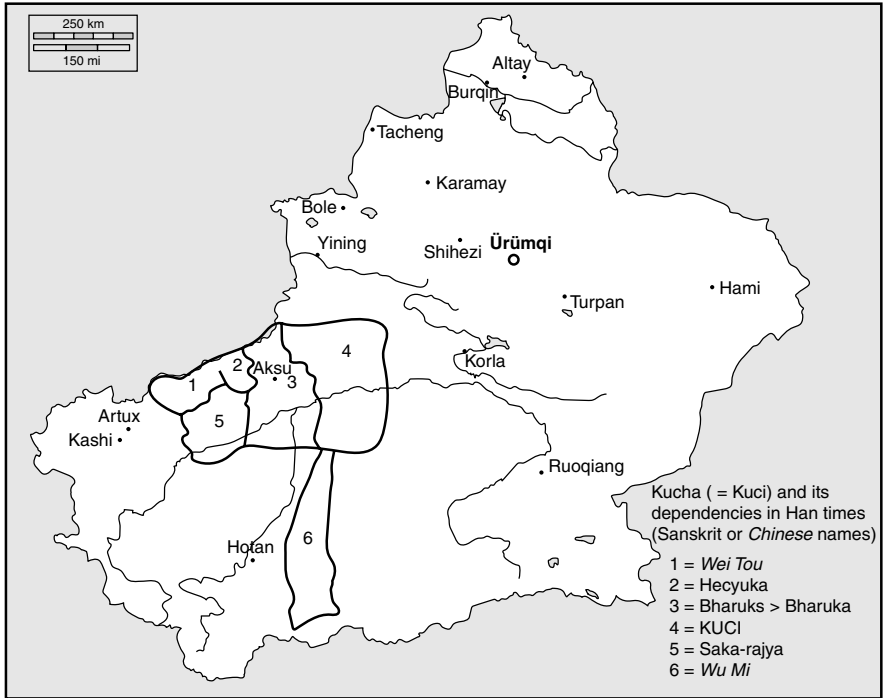
*Douglas Q. Adams*

## INTRODUCTION

The speakers of the Tocharian languages enter history in the first century BC, when the Han Empire sent military expeditions into the Tarim Basin (what is now the southern half of the Chinese province of Xinjiang). Our linguistic records date only to the late fourth century AD in the case of Tocharian B and to the seventh century AD in the case of Tocharian A. The latest documents in the two languages are probably no later than the ninth century AD (the last dated Tocharian B texts are from the final decade of the eighth century AD). The age of our texts, somewhat older than has often been thought in the past, is confirmed by both paleography and radio-carbon dating. It is also significant, with respect at least to Tocharian B, that our texts cover a half millennium of time, time sufficient for us to see significant change. When it was thought that all the attested texts from Tocharian B were more or less contemporary, within a century or so of one another, the linguistic differences among the texts were reasonably enough taken as the result of there being regional dialects (Winter 1955). However, now it is quite certain that at least most of the differences are the result of the documents' different ages (Peyrot 2008).

Tocharian B was the native language of the Kingdom of Kucha (Tocharian B *Kuśi*, Uyghur *Kucha*, Sanskrit *Kuci*, Chinese *Qiuci*), and Tocharian A that of the Kingdom of Agni (Uyghur *Qarashahr*, Sanskrit *Agni*, Chinese *Yanqi*) to the east of Kucha. Both languages were also used, at least as liturgical languages and languages of learning, in the Turpan Basin to the northeast. Almost all of the Tocharian A texts and the majority of those written in Tocharian B are of Buddhist content. But there are non-religious texts as well: medical texts, economic texts of one sort or another (inventories, bills of sale, etc.), travel documents/passports, and graffiti. Rare are original literary works (though by far the longest text in Tocharian A is a dramatic narrative, the *Maitreyasamiti*, based on Indian Buddhist models but composed in Tocharian A). Missing entirely in both languages are historical, mythological, or legal texts.

The languages were called "Tocharian" by early investigators because of the presumed connection of the people who spoke them and the classical *Tokharoi* who lived in what is now Uzbekistan but were known to have been driven there by the Huns from Gansu in China (east of Xinjiang, between Tibet to the southwest and Mongolia to the northeast). Though often doubted, the connection is probably correct; we may note, for instance, that their Turkish-speaking neighbors referred to the speakers of Tocharian A as the *Tuyri*. If the connection is true, we can see that, at the very outset of their entering into Central Asian history, the Tocharian and *Tokharoi* would have extended from the western part of the Tarim Basin to the outer marches of the then Chinese Empire (see Adams 2000). We are not certain what the native speakers called the two languages, hence the neutral, but banal, designations "Tocharian A" and "Tocharian B."



MAP 10.1 KUCHA (= KUCI) AND ITS DEPENDENCIES IN HAN TIMES

To the west of Kucha, in the vassal Kingdom of Gumo (this is the Chinese name; in Sanskrit records it is *Bharuka*), was spoken a language but little different from that of Kucha according to early Chinese observers. And to the southeast of the Tarim Basin, in the Kingdom of Kroraina (Chinese Loulan), was apparently spoken a related language (“Tocharian C” if you will) known only through some loanwords and loan morphemes (e.g., Kroraina Prakrit *-mci*, an adjective suffix with the meaning ‘pertaining to whatever noun it’s derived from’ and Tocharian A *-ñci* ‘id.’) from it into the official Indian language (“Kroraina Prakrit”) of the state (Mallory & Mair 2000: 278–279). To the south and southwest (beyond Gumo) in the Tarim Basin were Iranian languages, specifically two Saka Iranian languages, Tumshuqese and Khotanese. Tumshuqese was well within the cultural and political sphere of Kucha (Tumshuq’s “saka-rāja” was a vassal of the Kuchean king), while Khotanese was the language of a powerful state in the southern part of the Tarim Basin. The linguistic situation of the Tarim Basin appears to have been quite stable during the first Christian millennium until the last century or so thereof when all of these languages were replaced by the Turkish Uyghur, originally spoken northeastward of the Tarim Basin in what is now western Mongolia.

### Characteristics of the Tocharian languages and Indo-European languages

Though quite “advanced” phonologically when compared to contemporary Indo-European languages (and Tocharian A is more “advanced” than Tocharian B), the

Tocharian languages are otherwise very much “age-appropriate” Indo-European languages. They preserve both nominal and verbal dual forms, and thus, for the middle of the first Christian millennium, are in this matched in conservatism only by Slavic and Baltic (admittedly we have no idea about pre-Albanian in this time period). They have a large array of nominal cases (again matched during the period of Tocharian attestation only by Slavic and Baltic), though in several cases the exponents of particular case functions are clearly innovative (i.e., the “secondary” cases, on which see below). Among nominal categories, only the effacement of the neuter (so that it is identical with the masculine in the singular and with the feminine in the plural) is a clear structural innovation, and its “retreat” is a development matched by contemporary Romance and Celtic languages.

The verbal system is also “old-fashioned,” though of the Proto-Indo-European triad of present, aorist, and perfect, it has lost the finite forms of the perfect, preserving only its participle (combined with remnants of the aorist participle into a general “preterit” participle”). Of contemporary Indo-European verbal systems, it reminds one most of the Greek system and, to a somewhat smaller degree, the Romance system. Its core vocabulary is pretty much composed of straightforward inheritances from Proto-Indo-European (cf. the words for ‘father’, ‘mother’, ‘brother’, ‘sister’, which are (in Tocharian B/Tocharian A order) *pācer/pācar*, *mācer/mācar*, *procer/pracer*, *ṣer/ṣar*, which reflect pretty regularly PIE \*ph<sub>2</sub>tēr, \*meh<sub>2</sub>tēr, b<sup>h</sup>reh<sub>2</sub>tēr, \*swesōr). However, that core vocabulary is supplemented by a large number of borrowings from Sanskrit for Buddhist concepts and ideas.

What makes the Tocharian languages seem so far from the Indo-European mainstream is their pronunciation, their “innovative phonology.” Between the breakup of Proto-Indo-European and our first attestation of Tocharian, there were some major sound shifts. The PIE plain velars, (e.g., \*k) had fallen together with the palatal velars (\*k<sup>h</sup>); Tocharian shares this development with all the Indo-European groups of the west (Germanic, Italic, Celtic, Greek). Particular to Tocharian was the merger of the three series of PIE stops: voiceless, voiced, and voiced aspirate. Thus, for instance, the bilabials, \*p, \*b, \*b<sup>h</sup>, had all merged into a single Proto-Tocharian voiceless stop (e.g., \*p). By these two changes the fifteen PIE stops had become only four (\*p, \*t, \*k, \*k<sup>w</sup>) in Proto-Tocharian. On the other hand, pre- or Proto-Tocharian underwent a very thorough-going shift from unpalatalized consonants before front vowels to palatalized, so that \*t, \*k, and \*k<sup>w</sup> appeared in Proto-Tocharian as \*c, \*č, and (also) \*č in that environment (in attested Tocharian, both A and B, Proto-Tocharian \*č had become ś). Similarly, \*s, \*n, and \*l were palatalized to \*š, \*ñ, and \*λ (written as <ly>). Exceptionally, dentals had become \*ts when before \*y and PIE /d/ had disappeared (via *ǵ*?) or become affricated (ultimately to become <ts>) under certain circumstances difficult to specify. The PIE vowels also underwent major changes that are perhaps most easily grasped in list form; see Table 10.1 (the closing apostrophe, ’, indicates that the preceding consonant had been palatalized).

Taken together, these changes, to consonants and vowels, may make the PIE antecedents of a Tocharian form momentarily (or not so momentarily) opaque, e.g., PIE \*d<sup>h</sup>ewk- ‘hide’ > Toch. AB *cuk-*, PIE \*h<sub>1</sub>ekwos ‘horse’ > Toch. B *yakwe* [A *yuk*], PIE \*widwo- ‘wise’ > Toch. B *uwe*, PIE \*pode ‘(two) feet’ > Toch. B *pai(ne)* [A *pe(m)*]. The merger of the three orders of PIE stops (voiceless, voiced, and voiced aspirated) makes Tocharian etymologies more difficult to do and feel certain about than those in other major Indo-European groups because, say, a Tocharian verbal root *tāk-* might reflect a PIE initial \*t-, \*d-, or \*d<sup>h</sup>- and a final \*-k-, \*-k̄-, \*-g-, \*-ḡ-, \*-g<sup>h</sup>-, or \*-ḡ<sup>h</sup>-.

TABLE 10.1 THE DEVELOPMENT OF PIE VOWELS IN TOCHARIAN

PIE	Toch. B/Toch. A
*i	ʾä/ʾā (but <i>ä</i> in the neighborhood of PIE *s or *w; ʾi/ʾi when part of a living ablaut series)
*ī	ʾi/ʾi
*e	ʾä/ʾā
*e	ä/ā (automatic “reduced” vowel inserted to break up difficult consonant clusters)
*ē	ʾe/ʾa
*a	ā/ā
*H	ā/ā (*H indicates any vocalized laryngeal, vocalized only when between two obstruents)
*ā	ā/ā
*o	e/a (but <i>ä</i> when unstressed before word-final *m, *n, *y, *w, and *r)
*ō	ā/ā
*u	ä/ā (later than the change of *i/e to ʾä; u/u when part of a living ablaut series)
*ū	o/u
*ey	ʾi/ʾi
*oy	ei (later ai)/e
*ew	ʾu/ʾu
*ow	eu (later au)/o
*ay	ai/e
*aw	au/o

### Place within Indo-European

It is universally assumed that the pre-Anatolians were the first Indo-European subgroup to separate themselves from the rest, and it is widely assumed that the pre-Tocharians were the second subgroup to separate themselves. Both assumptions are probably correct. But the gap between the time of the pre-Anatolians’ separation and that of the pre-Tocharians’ was much longer than the gap between the pre-Tocharians’ going their own way and the time of the more general breakup of the remaining PIE speech group. Just one example of the difference in the relationship of Anatolian and Tocharian with the rest of the Indo-European speech community is the complete lack of both subjunctive and optative in Anatolian but the presence of both in Tocharian, albeit in a clearly less developed form (cf. VERBS *infra*) than shown by other Indo-European groups.

While the Proto-Indo-European that pre-Tocharian separated from cannot have been completely uniform, Tocharian does not show a close resemblance to any other Indo-European branch. Thus, the exact location where, dialectologically speaking, that separation took place is not easy to determine. However, the thoroughness with which Tocharian nouns and adjectives of whatever declension have become mixed with *n*-stem forms strongly suggests that at some point Tocharian had developed productive *n*-stem by-forms to other declensional types and that these *n*-stem forms had a definitizing value that contrasted with a non-definite value of the other kind of declensions. This development is seen in an even fuller form in Germanic (so also apparently in the history of Albanian – Matzinger 2006: 200). So also do some very specific lexical comparisons point in the direction of “northeastern Indo-European” (i.e., Germanic and Balto-Slavic) such as Germanic \*deug- ‘hide’ (cf. Old English *deogol* ‘secret, hidden’) and Toch. *tuk-* ‘hide’, both from a latish PIE \*d<sup>h</sup>ewk-, a metathesized variant of the more widely distributed

\*kewd<sup>h</sup>- (cf. Germanic \*hūd- ‘hide’, Greek κεύθε/ο- ‘hide’, and Iranian cognates), or Toch. B *proksa* (pl.) ‘millet’ beside Slavic *proso* (sg.) ‘millet’ (< PIE \*prokso/eh<sub>2</sub>-). The Toch. B ablative ending *-mem*, immediately from PIE \*-mons, points in the same direction, in that it shows an affinity with the plural oblique cases in *-m-* in Germanic and Balto-Slavic (compare particularly Old Prussian *-mans*; note that the \*-ns of Proto-Tocharian and Old Prussian are independent innovations on a more original \*-s) as opposed to the \*-b<sup>h</sup>- that characterizes these cases in other Indo-European languages (such as Italic, Celtic, Greek, Indo-Iranian, and perhaps Albanian). Even more specific is the preservation of the reduplicating syllable, PIE \*Ce- (where C = the reduplicated word-initial consonant), that is the mark of the PIE perfect as such, with a morphologically restricted retention of *Ce-* in both Gothic and Tocharian and not as \*\*Ci- (in Germanic) and \*\*C’ä- (in Tocharian) as the sound-laws of Germanic and Tocharian would otherwise require.

The dating of the breakup of that Proto-Indo-European unity that remained after the pre-Anatolians had departed is very much a matter of discussion. However, a date of 4,000 BC (± 500 years or so) is likely. As stated above, the pre-Tocharians probably left relatively early in the dissolution of the Proto-Indo-European speech community. How they got from the northern rim of the Proto-Indo-European-speaking world, adjacent to the pre-Slavs and pre-Germanics, and presumably somewhere on the Russo-Ukrainian steppe, to the Tarim Basin remains archeologically opaque. They are often thought to have been members of the Afanasevo culture of south central Siberia (so, for instance, Mallory & Mair 2000). That is not impossible. But any archeological traces of an extension of Afanasevo southward towards the Tarim Basin are faint. The timing would be problematic as well. The Afanasevo culture begins about 3,500 BC (lasting until about 2,500 BC), and if the pre-Tocharians were its bearers, there would have been precious little time for any interaction with early Indo-Iranian as described in the next paragraph. More importantly, perhaps, there are no traces of cereal agriculture associated with the Afanasevo culture, and cereal agriculture clearly had a continuous history among Tocharian speakers, from attested Tocharian back to the Proto-Indo-European homeland (witness *proksa* [pl.] ‘millet’, Proto-Slavic \*proso ‘millet’). Another possibility is that they were somehow participants in the Andronovo culture (or cultures) of northern Kazakhstan and adjacent southern Siberia (in various forms dating from 2,000–900 BC). Among the Andronovos some agriculture was always present. The graves of the earliest culture found in the Tarim Basin, the Qäwrighul culture (beginning about 1,800 BC; presumably the culture of the pre- or Proto-Tocharians), unhelpfully show some similarities with both their Afanasevo and Andronovo counterparts (Mallory & Adams 1997, s.vv. *Afanasevo*, *Andronovo*, *Qäwrighul*; Mallory & Mair 2000).

The pre-Tocharians and the pre-Indo-Iranians were all on the eastern frontier of the Proto-Indo-European world at one time or another. They would certainly have run into one another if the pre-Tocharians had spent an appreciable amount of time as members of the Andronovo culture. It is worth noting that there is no particular evidence for any early, i.e., PIE-era, association of Tocharian with the other eastern Indo-European languages, the Indo-Iranian languages. Certainly, as mentioned already above, Tocharian is a clear *centum*-language (i.e., the PIE palatal velars, \*k̑, \*g̑, and \*ǵ<sup>h</sup>, remain unpalatalized and fall together with PIE dorsovelars \*k, \*g, and \*g<sup>h</sup>), just as Indo-Iranian languages are clearly *satəm*-languages (where PIE \*k̑, \*g̑, and \*ǵ<sup>h</sup> did undergo palatalization and remain distinct from the dorsovelars \*k, \*g, and \*g<sup>h</sup>). There is only one lexical item clearly and certainly shared by just Tocharian, Iranian, and Indic: B *mašce* ‘fist’, Avestan *mušti-* ‘id.’, Sanskrit *muṣṭi-* ‘id.’ All go back unproblematically to PIE \*mustéy-. A second example is far less certain (because of the inexactness of the semantics): B *welke* (part of a plant),

Avestan *varāka-* ‘leaf’, Sanskrit *valkā-* ‘bark’. Any random Indo-European threesome might, quite accidentally as it were, share exclusive ownership of two words.

Interestingly enough, however, given the lack of shared items of certain Proto-Indo-European age, there is a small set of really quite ancient borrowings from Iranian to Tocharian and, in one case, from Indic to Tocharian. Examples (all from Tocharian B) are *tsain* (pl. *tsainwa*) ‘arrow’ (< \*Proto-Iranian \*dzainu-, cf. Avestan *zaēna-* ‘armor, weapon’, Armenian *zēn* (a *u*-stem) ‘id.’, also borrowed from Iranian), *tsaiññe* ‘ornament’ (< Proto-Iranian \*dzai- ‘to ornament’), *waipecce* ‘possession’ (< Proto-Iranian \*hwaipatya-, cf. Avestan *hvaēpaiθya-* ‘ones’ own’), *šem* ‘axle’ (< Proto-Iranian \*axšám, cf. Gathic Avestan *aša-*), *išcem* ‘brick, tile’ (i.e., ‘fired clay’) (< Proto-Iranian \*ištýám, cf. Late Avestan *ištiia-* [n.]), *ekšinek* ‘dove’ (< Proto-Iranian \*akšínaka-, cf. Ossetic *æxšínæg*), and \*wāstare ‘camel’, unattested but cf. *wāstarye* ‘pertaining to a camel’ (< Proto-Iranian \*ustrá- ‘camel’, cf. Avestan *uštra-*). These borrowings are significant because they show an early stage of Iranian phonological development: preservation of the affricate \*-dz- (in attested Iranian always \*-z-), and of the cluster \*-ty- (rather than later \*-θy-); word-final stress; or even, once, a form that had not undergone the RUKI rule, i.e., \*ustrá-, not \*uštrá-. The Iranian we see through the window of these early Tocharian borrowings noticeably antedates the Avestan of Zarathuštra. Since Zarathuštra is commonly, though not universally, dated to about 1,000 BC, that puts these early Iranian-Tocharian contacts well into the second millennium BC. Whether these contacts took place inside the Tarim Basin or outside cannot be said.

The singular example of a demonstrably early borrowing from Indic is *kercao* ‘ass’ (= Sanskrit *gardabhá-* ‘id.’ < pre-Indo-Aryan \*gordeb<sup>h</sup>ó-). It is particularly significant, however, because that borrowing must have taken place before the merger of PIE \*e, \*a, and \*o in Indo-Aryan and so even earlier than the early borrowings from Iranian where that merger had already taken place. This merger had already been accomplished by 1500 BC as seen in Indo-Aryan loanwords borrowed into Mitanni in eastern Asia Minor. In this case we can be certain the borrowing must have occurred before the pre-Tocharians entered into the Tarim Basin or the pre-Indo-Aryans left Central Asia for India (the latter event being, on most accounts, sometime between 2,000 BC and 1,500 BC).

Out of all these considerations a picture, both fuzzy and speculative, begins to emerge of the pre-Tocharians leaving the northern frontier of the Proto-Indo-European world and moving east about 4,500 BC or so to the southern Urals-North Kazakhstan steppes. There they had limited contact with both Proto-Indo-Aryans (first) and Proto-Iranians (second), already dialectally distinguished from one another, in the period, say, 2,500–2,000 BC. After that contact the Proto-Tocharians moved further east, into the Tarim Basin, while the Proto-Indo-Aryans and Proto-Iranians, in that chronological order, moved further and further south (with one arm of the Proto-Iranians, the Khotanese and Tumshuqese, also eventually entering the Tarim Basin and settling the southwestern portion of it). Such a scenario would allow the pre-Tocharians to be in the Tarim Basin as the bearers of the Qäwrighul culture of the eastern Tarim Basin (beginning about 1,800 BC), a culture known almost exclusively from its burial sites and spectacular mummies. The area occupied by the Qäwrighul culture was later the nucleus of the historical Kingdom of Kroraina (Chinese Loulan), where we find the remnants of “Tocharian C.”

Contact with Iranian was maintained, or renewed (and now clearly with the Tocharians in the Tarim Basin), in the middle of the first millennium BC or thereafter. We find such borrowings with later Iranian phonology as *newiya* ‘(a kind of) canal’ (< Proto-Iranian \*nāwiya- ‘navigable’), *murye* ‘(a kind of) canal’, *ärtte* ‘(feeder) canal’, *peri* ‘debt’ (< Proto-Iranian \*parya-), *sāte* ‘rich’ (< Proto-Iranian \*čyāta-), which suggest a very different

cultural level of inter-language contact. Somewhat later, the advent of Buddhism in the Tarim Basin brought with it a host of religious/cultural terminology, either as borrowings or as calques, from Sanskrit and, to a lesser extent, from Iranian.

## PHONOLOGY

The two Tocharian languages have exactly the same phonemic inventories.

**TABLE 10.2 TOCHARIAN CONSONANTS**

	Bilabial	Apico-Dental	Apico-Alveolar	Lamino-Palatal	Dorso-Velars
stops, affricates	<i>p</i>	<i>t</i>	<i>ts</i>	<i>c</i>	<i>k</i>
continuants			<i>s</i>	<i>ś</i>	
nasal stops	<i>m</i>	<i>n</i>		<i>ñ</i>	
laterals		<i>l</i>		ɹ <ly>	
trill		<i>r</i>			
approximants	( <i>w</i> )			<i>y</i>	<i>w</i>

**TABLE 10.3 TOCHARIAN VOWELS**

<i>i</i>		<i>u</i>
	<i>ä</i>	
<i>e</i>	<i>ā</i>	<i>o</i>

The symbol <ä> represents something like IPA [i] and <ā> something like IPA [a].

**STRESS:** It is possible that both Tocharian languages have contrastive stress, but only Tocharian B shows it in ways visible to us, as described below (under the heading “Morphophonology”). This contrastive stress may result in minimal pairs, e.g., *tākam* ‘we will be’ (</tākām/ and *takām* ‘we were’ (</tākām/)).

**WRITING SYSTEM:** Though neither Tocharian language has been spoken for a thousand years or so, we can be quite confident that our knowledge of the phonology of both languages is accurate and substantially complete. That is so because they are written in a form of Brahmi, the Indian alphabet whose variants have been used to write various Indian and Indo-Chinese languages since the last centuries before the Christian era. The phonetic values of the signs of the alphabet are well known. Unique to the Tocharian variant of this writing system are that so-called Fremdzeichen (‘foreign signs’) created by those who adapted Brahmi to Tocharian to signal the vowel /ä/, unknown to Sanskrit and other Indic languages.

**MORPHOPHONOLOGY:** In Tocharian there are three major, productive, morphophonological processes, which shape many Tocharian paradigms: palatalization, umlaut, and ablaut. In addition, there are widespread, purely phonological, processes that account for the distribution of [ā], [a], and [ä] in Tocharian A and Tocharian B. We will take up these latter processes first.

**DISTRIBUTION OF [Ā], [A], AND [Ä] IN TOCHARIAN B:** In Classical and later Tocharian B the central vowel /ä/ is strengthened in stressed syllables to [a]. In unstressed closed

syllables it remains as [ä], while in unstressed open syllables it disappeared completely (e.g., /stämó-/ ‘having stood’ > *stmó-*). Similarly, in stressed syllables /i/ and /u/ may be reflected in the spelling as <ɪ> and <u> respectively (i.e., the long vowel signs are used to show stressed syllables). On the other hand, in stressed position /ā/ remains unchanged, but, conversely, in an unstressed syllable /ā/ appears as [a]. Both the strengthening of /ā/ (and /i/ and /u/) and the “destrengthening” of /ā/ are shown orthographically in Classical and Late Tocharian B. To give but one example: in Archaic Tocharian B the third person singular of ‘(s)he released’ would have been *cārkā*, but in Classical and Late Tocharian B it was *carika*. It is from the strengthening of /ā/, the destrengthening of /ā/, and the loss of unstressed /ā/ in open syllables that we derive the majority of our information about Tocharian B stress, since stress itself is never marked orthographically.

DISTRIBUTION OF [Ā], [A], AND [ǻ] IN TOCHARIAN A: Superficially resembling in some ways the interchange of [ā], [a], and [ǻ] in Tocharian B is a similar interchange of the surface representatives of /ā/ in Tocharian A. Though considerably disturbed by analogical restorations of one sort or another, the basic historical pattern is like this: if, in a word of two syllables or more, the first syllable contained /ā/, /e/, or /o/ and the second syllable contained /ā/, then the second syllable’s /ā/ shows up on the surface (1) as [a] if the second syllable is word final and closed, (2) as [ǻ] if the second syllable is not the final syllable of the word but is a closed syllable, and (3) as zero if the second syllable is not the final syllable of the word and is an open syllable.

### Condition 1:

*kotnaš* ‘(s)he cuts off’ from /*kotnāš*/ (cf. Toch. B *kautanaṃ* < PToch. \**kāutnā-*, or, within Tocharian A when the first syllable of the word is “light,” A *kārsnās* ‘(s)he knows’)

*kropat* ‘(s)he gathered’ (cf. Toch. B *kraupāte* < PToch. \**krāupāte* or, within Tocharian A, *kālpāt* ‘(s)he achieved’)

*šāmaṃ* ‘monk’ (B *šamāne* < PToch. \**šāmāne*)

*āknats* ‘foolish’ (B *aknātsa* < PToch. \**āknātsā*)

### Condition 2:

*tāpākyāñ* ‘mirrors’ (cf. *tāpaki* [< \**tāpaky*] ‘mirror’)

*kākārpau* ‘having descended’ (B *kakārpau* < PToch. \**kākārpā-*)

### Condition 3:

*kropte* ‘thou has gathered’ (cf. B *kraupātai* < PToch. \**krāupātāi*, or, with Tocharian A, *kālpāte* ‘thou hast achieved’)

*krophlune* ‘gathering’ (cf. B *kraupālñe* or, within Tocharian A, *kālpālune* ‘achievement’)

*kākmu* ‘having brought’ (cf. B *kakāmau* < PToch. \**kākāmā-*)

In addition, as in Tocharian B, /ā/ disappears in open syllables, even from those syllables that, on the basis of Tocharian B evidence, were stressed.

PALATALIZATION: Palatalization (as already outlined above) developed when a susceptible root-initial consonant or consonant cluster preceded a PIE \*-e- or \*-i-, whether short or long (short \*-i- does not palatalize if it is further in the environment of \*-w- or \*-s-).



Thus, we have  $k > ś$  (via \*č),  $t > c$ ,  $ts > ś$  (Tocharian A only),  $n > ñ$ ,  $l > ly$ ,  $w > y$  (Tocharian B only),  $s > ś$ ,  $sk > śś$ ,  $st > śc$ ,  $sp > śp$ ;  $p$ ,  $m$ ,  $r$  are not palatalizable in either language, and neither are any of the consonants that can be the result of palatalization ( $c$ ,  $ñ$ ,  $y$ , etc.). However, in Tocharian B certain preterits to marked causative roots exceptionally allow “palatalization” of most of the non-palatalizable consonants. Thus, we find the “special” or “secondary” palatalization:  $k > ky$ ,  $ñ > ñy$ ,  $p > py$ ,  $m > my$ ,  $ts > tsy$ . Non-causative preterits show only “primary” palatalization, e.g., Toch. B *lyāka* ‘he saw’ and Toch. A (imperfect) *lyāk* ‘he was seeing’.

Root-initial palatalization is strongly favored by causativity. Almost all preterits to morphological causatives in Tocharian B are preterits that uniformly show palatalization, regular and “special,” where possible. Causative presents or subjunctive forms with the suffix *-āsk-* sometimes, but not commonly, also show palatalization: B *śarsāsk’ā/e-* [A *śarsās’ā/a-*] ‘teach’, causative to *kārs-* ‘know’ (non-causative present *kārsnā-* in both languages), B *śatkāsk-* (beside *katkāsk-*) ‘send over’ to *kāt-* ‘cross over’ (both with preterit *śātka*), B *śalāsk’ā/e-* ‘to throw’ (but A *sālās’ā/a-*) to *sāl-* ‘fly’, B *śpantāsk’ā/e-* ‘convince’ to *spānt-* ‘trust’, B *śparkāsk’ā/e-* ‘destroy’ to *spār-* ‘perish’, and *śpartāsk’ā/e-* (~ *śparttask’ā/e-* ~ *spārttask’ā/e-*) ‘turn (trans.)’ to *spārt-* ‘turn (intr.)’. Note, with “secondary” palatalization we have B *pyutkāsk’ā/e-* [A *pyutkās’ā/a-*] ‘come into existence’. But palatalization may also mark the present as opposed to the subjunctive, e.g., *ceśām/cekem* ‘touch(es)’ but *tekām/takām\** ‘will touch’ (A present *cāk’ā/a-*, subjunctive *tākā-*) or B *cepiy(e)-* ‘step forth, appear’ but *tāpā-* ‘will appear’.

UMLAUT: Both Tocharian languages have undergone a change whereby an *\*-ā-* in a following syllable has changed a preceding *\*-(‘)e-* to *\*-(‘)ā-*, a phonological process we can call “*ā*-umlaut.” The change works somewhat differently in Tocharian A and B, and that difference suggests the change was only in process at the time, whenever that was, that the two varieties of Proto-Tocharian that were to become Tocharian A and Tocharian B were losing contact with one another. In Tocharian B any *\*-(‘)e-* became *\*-(‘)ā-* when another *\*-ā-* followed in the next syllable. In Tocharian A the change occurred only when the *\*-(‘)e-* was, on the basis of the Tocharian B evidence, unstressed. There is a common subjunctive formation that has the root-vowel *\*-e-* (< PIE *\*-o-*) in the active singular and *\*-ä-* elsewhere. When the thematic vowel, the morpheme joining root and person-number endings, is *-ä-*, the root-vowels remain unchanged, thus B *tekām/takām* (< *\*tékān/tākān*) ‘(s)he will touch/they will touch’. By normal change Proto-Tocharian *\*-e-* becomes Tocharian A *-a-*, but otherwise there is no change in the latter language either. When the thematic vowel is *\*-ā-*, however, Tocharian B shows the effect of *ā*-umlaut, e.g., *tārkaṃ/tarkaṃ* (< *\*térkān/tárkān*) ‘(s)he will release/they will release’. In Tocharian A the result is *tarkaṣ/tärkeñc* with the regular change of *\*-e-* to *-a-* but with no *ā*-umlaut to *\*\*tārkaṣ*. However, PToch. *\*lyekā-* ‘saw’ gives *lyākā-* in both languages (3 sg. B *lyāka*, A *lyāk*). The effect of *ā*-umlaut is analogically effaced in some situations. Thus, the plural of B *pīle* ‘wound’ is *pīlenta*, with preserved *-e-*, and not the *\*\*pīlānta* required by *ā*-umlaut. But in verbal morphology *ā*-umlaut is quite regular. Similar to *ā*-umlaut, but much rarer, is *o*-umlaut, e.g., /*sesoyu-* ‘satisfied’ > B *soṣoyu-*.

There is another kind of vowel affection that is to be seen in Tocharian B: in many situations in the Tocharian B verb where we expect *ā . . . e* or *e . . . ā* what we find is *o . . . o* instead. Thus, the imperative of *āks-* ‘speak, announce’ should be /*pe-ākse-* underlyingly but appears in the surface structure as *pokse* instead. Likewise, Class IV presents, parallel in every way to Class III presents save that Class IV’s have the root-vowel *-ā-* instead of

Class III's *-ä-*, have the shape, for example, B *kloyotär* '(s)he falls' rather than *\*\*kläyétär* (compare the Class III B *pälkétär* 'burns' (intr.)).

ABLAUT: Both Tocharian languages show abundant remains of PIE ablaut, particularly in verbal paradigms. Ablaut is not found paradigm-internally in present formations but is relatively common in both the subjunctive and preterit. As described above, many show a pattern of ablaut wherein the active singular is distinguished by one vowel (*-e-* [A *-a-*] or *-ä-*, while the rest of the paradigm has *-ä-*). In the preterit there are two loci for inter-paradigmatic ablaut: (1) in certain *ā*-preterits ("Class I") and (2) in sigmatic preterits ("Class III").

For Class I preterits with ablaut, the usual ablaut situation is an *-ä-* in the active with preceding palatalization and *-ä-* with no palatalization in the medio-passive and preterit participle, e.g., B *cārka* '(s)he released', B *cärkäre* 'they released' vs. B *tärkä(n)te* '(s)he was/they were released' (and preterit participle *tärkäu*). In A the situation is different, and demonstrably older. In the active singular there is an *-ä-* (as in B), *cärk* '(s)he released', but the preterit (dual and plural) have *-a-*, e.g., *tarkar* '(s)he released'. The medio-passive (*tärkä(n)t*) and participle (*tärko*) have non-palatalizing *-ä-*, as in Tocharian B. Tocharian B has one example of the older, Tocharian A, situation, where we find underlying *-ä-* with preceding palatalization in the active singular, *-ä-* (< underlying (Toch. B) *-e-* by *ā*-umlaut) in the non-singular active, and *-ä-* with no palatalization elsewhere: *ścāma* [3 act. sg.], *stāmáis* [3 act. dual], and (once) *stamāre* [3 act. pl.] from *stäm-* 'stand'. The active singular, then, represents an old PIE \*e-grade and the medio-passive an old PIE zero-grade. In Indo-European terms the more archaic type would appear to reflect an (active singular) \*e-grade, (active plural) \*o-grade, and (medio-passive) zero-grade. A generally accepted explanation for this ablaut system remains to be found.

Class III or sigmatic preterits offer two ablaut (or rather palatalization) patterns. In the first we have Tocharian B *-e-* in the active, usually in the medio-passive, and the preterit participle – all with no palatalization. In some verbs of this group the medio-passive (MP) and preterit participle (PP) have Tocharian B *-ä-*. An example with constant *e-* is *täm-* 'be born' (*temtsamai* [1 MP sg.], *temtsate* [3 MP sg.], *temsamte* [1 MP pl.], *temsamte* [3 MP pl.]; PP *tetemu*). Examples with an interchange of *-e-* and *-ä-* are *pärk-* 'ask a question' (*preksa* [3 act. sg.], *parkante* [3 MP pl.]; PP *peparku*; A *prakäs* [3 act. sg.], *präksāt* [3 MP sg.], PP *papärku*) and *räk-* 'extend [one's hand]' (*reksa* [3 act. sg.; cf. Latin *rēxit*], *raksamai* [1 MP sg.], *raksate* [3 MP sg.]; PP *reraku\**; A *rakäs* [3 act. Sg.], PP *rarku*). In the second subgroup we find the active with palatalization + *-e-* and the medio-passive and preterit participle with *-ä-*. A good example is the causative paradigm of *pälk-* 'burn, torture' (*pelykwa* [1 act. sg.], *pelyksa* [3 act. sg.], *palyksatai* [2 MP sg.]; PP *pepalyku\**). In Indo-European terms the first, non-palatalizing, subtype would appear to reflect the (active) \*o-grade and (medio-passive) zero-grade, while the second, palatalizing, subtype would appear to reflect the (active) \*ē-grade and (medio-passive) \*e-grade. Tocharian A shows the same two patterns (with *-a-* of course rather than *-e-*) but differs greatly from Tocharian B in the lexical distribution of the two types.

## MORPHOLOGY

The Tocharian languages are typically Indo-European in their complex morphology, using a rich set of inflections in both nouns and verbs.

**NOUNS:** A Tocharian noun belongs to one of three **GENDER CLASSES**, masculine, feminine, and neuter. The last class is also referred to as alternating, as nouns in this class look identical to masculine ones in the singular and feminine ones in the plural. Nouns inflect for **NUMBER:** singular, dual, and plural, and **CASE.** (For a discussion of the Tocharian dual and the supposed “paral” (for natural pairs), see Winter 1962.) The various cases mark the nouns’ syntactic role in a sentence. In Tocharian B there are nine cases: nominative, accusative, genitive, vocative (only distinct from the nominative in the singular), ablative, causal (only in the singular), comitative, locative, and perlativ. The first four are the so-called primary cases, and the last five are secondary cases.

The primary cases are developments of PIE case endings; have different forms for singular, dual, and plural; are always marked on the noun; and, in many nouns, cause a shift in stress. For example, the nominative and accusative of *yákwe* ‘horse’ have stress on the first syllable, but the genitive is *yákwéntse* with stress on the second. (Evidence for accent-shifting can only come from Tocharian B.) This mobility of accent does not in general reflect anything of PIE age. By a variety of regular phonological and analogical changes, not always well understood, nominal stress came usually to be on the first (left-most) “strong” vowel (-e-, -o-, -ā- and diphthongs) and not on a “light” vowel (-i-, -u-, -ā-). Thus, an original PIE \*h<sub>1</sub>ékwo- ‘horse’ became PToch. \*yákwé-. Complicating the surface structure of end-stressed nouns, like \*yákwé-, at least in Tocharian B, is a later phonological development, Marggraf’s Law, that retracted stress from absolute final syllables, and thus the nominative-accusative singular \*yákwé > \*yákwē, but the genitive, \*yákwén(ā)se, remained unaffected.

Adjectives agree in case with their head nouns when the head nouns are in the primary cases. Secondary cases, on the other hand, are thought to derive from PIE postpositions that have become fused with the noun, have the same shape whatever the number of the noun they is attached to, occur only once on a series of nouns in the same case, and do not cause any accent shift. Adjectives do not agree with their head nouns when the head nouns are in any of the secondary cases. When modifying nouns in secondary cases, adjectives are uniformly in their accusative shapes. However, the distinction between primary and secondary cases is not absolutely sharp. In particular, the Toch. B ablative in *-mem* often causes accent shift, giving rise to the possibility that it was once a primary case ending that has become “secondarily a secondary case” if you will and the same is true of the Toch. A ablative ending *-s* (< PIE \*-ti, represented also by Hittite *-(a)z < \*(o)ti*).

A rather typical nominal paradigm would be *yákwe* ‘horse’ from PIE \*h<sub>1</sub>ékwo- (the Toch. A equivalent *yuk*, is given in italics in square brackets); see Table 10.4.

**TABLE 10.4 AN EXAMPLE OF TOCHARIAN NOMINAL DECLENSION**

	Singular	Dual	Plural
Nom.	yákwe [yuk]	yákwene [yukäm]	yákwi [yukañ]
Acc.	yákwe [yuk]	yákwene [yukäm]	yákwem [yukas]
Voc.	yákwa [yuk]	yákwene [yukäm]	yákwi [yukañ]
Gen.	yákwéntse [yukes]	yákwénaisän [yuknis]	yákwémts [yukasśi]
Abl.	yákwemem ~ yákwémem [yukas]	yákwenemem [yuknas]	yákwemmem [yukasäs]
All.	yákwešc [yukac]	yákwenešc [yuknac]	yákwemšc [yukasac]

Com.	yákwempa [yukaśśāl]	yákwenempa [yuknaśśāl]	yákwemmpa [yukasāśśāl]
Loc.	yákwene [yukam]	yákwenene [yuknam]	yákwemne [yukasam]
Perl.	yákwesa [yukā]	yákwenesa [yuknā]	yákwemsa [yukasā]
Instr.	[yukyō]		[yukasyō]

This paradigm is directly descended from PIE masculine *o*-stems. However, like virtually every nominal or adjectival paradigm in Tocharian, it shows some importations from PIE *n*-stems. Thus, the genitive singular and genitive plural are from \*-o-neso and \*-o-nesom respectively, and the entire dual shows an *-n-* from the same source.

The fusion of accusative plus following postposition that created the secondary cases was apparently only beginning in the last stages of Proto-Tocharian because the two languages show different results. For instance, the allative postposition must have been PToch. \*-cā (= Greek *-de?*), and the perlicative postposition must have been \*-ā. When added to the singular accusative we would have had \*yákwē-cā for the allative; when added to the plural, \*yákwens-cā. For the perlicative it would have been singular \*yákw(e)-ā and plural \*yákwens-ā. At some point in the history of Tocharian A the accusative plural ending was reduced from \*-ns to \*-s. The allative would then have been singular \*yákwā-cā and plural \*yákwās-cā. The latter form was rebuilt, by analogy to the singular as \*yákwās-acā, whence the actual forms *yukac* and *yukasac*. The perlicative followed the same process, and \*yákwā-ā and \*yákwās-ā gave actual *yukā* and *yukasā*. In pre-Tocharian B, however, word-final \*-ns gave \*-n (written *-m* in attested Tocharian). Original \*yákwē-cā and \*yákwens-cā would have given \*\*yákwēc and *yákwēnc* respectively. Original \*yákwē-ā and \*yákwens-ā would have given \*\*yákwā and \*yákwensā respectively (with a morphological division \*yákwen-sā because now the accusative plural was \*yákwen). In both cases the plural form was generalized in Tocharian B, thus singular \*yákwēscā and \*yákwēsā. Also generalized from the plural was the Toch. B ablative ending *-mem*. In Tocharian A it was the singular ablative ending, in *-s*, that was generalized to the dual and plural.

**ADJECTIVES:** Tocharian adjectives are morphologically much simpler than nouns. They have at most four case forms (nominative, accusative, genitive, and vocative). When they modify nouns in other cases (e.g., ablative, instrumental) the accusative form is used, and, more often than not, nouns in the genitive are also modified by adjectives in the accusative. In form the most common kinds of adjectives are medleys of PIE *o*-stem, *yo*-stem, and *n*-stem adjectives. As in the case of nouns, the many *n*-stem forms that show up in many places in the Tocharian adjective reflect the pre-Tocharian situation whereby nouns and adjectives had productive *n*-stem derivatives that were individualizing and definite when compared to the non-*n*-stems that underlay them. With regard to adjectives, the situation must have been very similar to that of Proto-Germanic with the latter's distinction of definite and indefinite adjectives, the former being built on *n*-stem antecedents and the latter not. In Tocharian the two types of adjectives were fused. The fusion seems to have been largely post-Proto-Tocharian because the mechanics of that fusion are often quite different in the two languages. In Table 10.5 we give examples of an inherited (PIE) *o*-stem (*ñuwe* 'new' [Toch. B only; Toch. attests only *ñu* (m. sg.) and *ñwam* (f. pl.)]), *yo*-stem (*ñākciye* [*ñākci*] 'divine'), and two inherited *n*-stems (*krošce* [*k<sub>u</sub>raś*] 'cold' and *klyomo* [*klyom*] 'noble').

TABLE 10.5 SOME EXAMPLES OF TOCHARIAN ADJECTIVAL DECLENSION

m. sg. nom.	ñuwe	ñäkiye [ñäki]	krošce [kuraš]	klyomo [klyom]
acc.	ñuwe	ñäkiye [ñäki(m)]	kroścäm [kroššäm]	klyomom [klyomänt]
gen.	ñwepi	ñäkiyepi [ñäkināp]		klyomopi [klyomäntāp]
m. pl. nom.	ñuwi	ñäki [ñäkiñi]	krošci [krošše]	klyomoiñ [klyomäš]
acc.	ñuwem	ñäkiyem [ñäkinäs]	kroścäm [kroššes]	klyomom [klyomäncäs]
gen.	ñweṃts	ñäkiyemts [ñäkinäšši]		klyomomts [klyomäncäšši]
f. sg. nom.	ñuwa	ñäkiya [ñäki(m)]	[krošši]	klyomña [klyomiñ]
acc.	ñuwai	ñäkiyai [ñäkyām]	[kroššäm]	klyomñai [klyomiñām]
gen.	ñuwai	ñäkiyai [ñäkciñe]		klyomñai [klyomiñe]
f. pl. nom.	ñwona	ñäkiyana [ñäkyāñ]	kroścana [kroššāñ]	klyomñana [klyomiñāñ]
	ñwona	ñäkiyana [ñäkyās]	kroścana [kroššās]	klyomñana [klyomiñās]
	ñwonaṃts	ñäkiyanaṃts [ñäkinäšši]		

Similar in declension are the various productive adjectives in *-šše* [-ši] ‘pertaining to X’, *-ññe* [-ñi] ‘ibid.’, *-tse* [-ts] ‘having X’, and *-lle* [-l] ‘X-able’. Adjectives in *-lle* [-l], in both languages, and *-tse*, in Tocharian B only, show palatalization in the masculine singular (except the masculine singular nominative), masculine plural, and feminine singular (e.g., Toch. B (m. sg.) *-lle*, *-lye*, *-lyepi*, m. pl. *-lyi*, *-lyem*, f. sg. *-lya*, *-lyai*, f. pl. *-llana*, *-llana*). The palatalization is ultimately from the *n*-stems (in the masculine) and *yā*-stems (in the feminine). Tocharian has, in addition, abundant traces of participles in *\*-nt-* or *\*-us-* (both of which have been worked into the Toch. A paradigm for *klyom*) and the denominative *\*-went-* ‘having X’.

DEICTICS AND THIRD PERSON PRONOUNS: The different Tocharian deictics (i.e., the equivalents of ‘the’, ‘this’, and ‘that’), used also as third person pronouns, are based on a single basic paradigm to which different particles are attached. In Tocharian B the basic deictic/pronoun, usually used as a simple anaphoric, is masculine *se* (nom. sg.), *ce* (acc. sg.), *cey* (nom. pl.), *cem* (acc. pl.), feminine *sā* (nom. sg.), *tā* (acc. sg.), *toy* (nom./acc. pl.), neuter *te* (nom./acc. sg.). Only here in Tocharian is there a neuter, used predicatively in reference to whole clauses, or concepts. Except for the word-initial palatalization in some forms, the Tocharian B forms of the first deictic match very well with their counterparts in Greek (*se*, *ce*, *cey*, *sā*, *tā*, *te* = Greek *ὁ*, *τόν*, [Doric] *τοί*, *ση*, *τήν*, *τό*).

The second deictic, more or less equivalent to the English definite article, though used more sparingly, adds *-w* in Tocharian B and *-m* in Tocharian A. The third deictic, ‘this

one here', adds *-n* (written <ṃ>) in both languages. The fourth deictic, 'that there' (only in Tocharian B), adds *-m(p)*. As is commonly the case, the basic system is shared by both Tocharian A and B, but the details, and the inter-paradigmatic analogies, are the province of each language separately. The Tocharian B paradigms are underdifferentiated in the genitive and masculine plural, while Tocharian A has generally created unique forms for each "slot."

TABLE 10.6 DEICTIC PRONOUNS IN TOCHARIAN

	first		second		third		fourth
	B	A	B	A	B	A	B
m. sg. nom.	se	<i>sās</i>	su	<i>sām</i>	seṃ	<i>saṃ</i>	samp
m. sg. acc.	ce	<i>caṣ</i>	cew	<i>cam</i>	ceṃ	<i>caṃ</i>	camp
m. sg. gen.	cwi	<i>caṣi</i>	cwi	<i>cami</i>	cwi	<i>cani</i>	cwimp
m. pl. nom.	cey	<i>ceṣ</i>	cey	<i>cem</i>	cey	<i>ceṃ</i>	caimp
m. pl. acc.	ceṃts	<i>cesās</i>	ceṃts	<i>cesām</i>	ceyna	<i>cesām</i>	*caimpa
m. pl. gen.	ceṃts	<i>cessi</i>	ceṃts	<i>cesmi</i>	ceynaṃts	<i>cesni</i>	cempaṃts
f. sg. nom.	sā	<i>sāṣ</i>	sā <sub>n</sub>	<i>sām</i>	sāṃ	<i>sām</i>	somp
f. sg. acc.	tā	<i>tāṣ</i>	tā <sub>n</sub>	<i>tām</i>	tāṃ	<i>tām</i>	tomp
f. sg. gen.	tāy		tāy	<i>temi</i>			
f. pl. nom.	toy	<i>toṣ</i>	toṃ	<i>tom</i>	toyna	* <i>toṃ</i>	toym
f. pl. acc.	toy	<i>tosās</i>	toṃ	<i>tosām</i>	toyna	<i>tosām</i>	toym
f. pl. gen.			toṃts	<i>tosmāśsi</i>	toynaṃts		
n. sg.	te	<i>tāṣ</i>	tu	<i>tām</i>	teṃ	<i>taṃ</i>	tamp
n. sg. gen.			tuntse	<i>tmis</i>			
du.			tai	<i>tim</i>			
du. gen.			tainaisāñ				

RELATIVE AND INTERROGATIVE PRONOUNS: The commonest relative and interrogative pronoun, one that does not distinguish gender (in both languages) nor number (in Tocharian B), but that does distinguish case (in both), is a combination of the first deictic with a preceding *kw-* to give B nom. *k<sub>n</sub>se* (A *kus*), acc. *k<sub>n</sub>ce* (A *kuc*). The genitive forms are both irregular and not reducible to a single Proto-Tocharian preform; we have B *ket(e)* and A *ke*. In Tocharian A, but not in B, the relative pronoun is distinguished from the interrogative by the addition of the particle *ne* to the former but not the latter (thus *kusne* and *kucne*). In Tocharian A a nominative plural of the relative pronoun is also known, *kucene* (no accusative plural is known, perhaps simply because of a lack of attestation). Both languages have a relative or interrogative adjective derived from a deictic prefixed by *in-* (in B) and *än-* (in A). In Tocharian B it is the second deictic that is so used; in Tocharian A, the third. Thus, we have B *intsu* (nom.), *iñcau* (acc.), etc. 'which' [A *äntsam*, *äñcam*]. More commonly, in this function in Tocharian B we have the second deictic prefixed by *mäk-*, i.e., *mäksu* (nom.), *mäkcau* (acc.), etc.

FIRST AND SECOND PERSON PRONOUNS: The array of personal pronouns in the two Tocharian languages is given in Table 10.7.

TABLE 10.7 FIRST AND SECOND PERSON PRONOUNS IN TOCHARIAN

	Toch. B sg.	Toch. A m. sg.	Toch. A f. sg.	Toch. B pl.	Toch. A pl.
nom.	ñäš	näš	ñuk	wes	was
acc.	ñäš	näš	ñuk	wes	was
gen.	ñi	ñi	nāñi	wesāñ	wasām
nom.	tuwe	tu		yes	yas
acc.	ci	cu		yes	yas
gen.	tañ	tñi		yesāñ	yasām

In addition there is a reflexive genitive B *šañ* [A *šñi*] and, attested only in Tocharian B, dual nominative-accusative *wene* ‘we too’ and *yene* ‘you too’.

All of the second person forms, and the dual and plural of the first, have direct PIE etymologies (*tuwe* < \*tuHóm [cf. Sanskrit *tuvám*], *wes* < \*wos [cf. Latin *uos*]) or are easily explained analogical developments of known PIE forms. Not so the first person singular. The initial *ñ-* may reflect a PIE oblique case stem \*mne-, but essentially everything else is unclear or downright weird. The *-š* of Tocharian B and the *-š* of Tocharian A are not reducible to a common denominator, and the origin of the typologically very unusual distinction in gender in Tocharian A is without any accepted explanation.

NUMBERS: Cardinal numbers are inflected for gender (‘one’, ‘two’, ‘three’, and ‘four’ in Tocharian A; ‘one’, ‘three’, and ‘four’ only in Tocharian B), number (‘one’ only), and case (a full panoply only for ‘one’, a single nominative-accusative form and a rarely attested genitive for the other numerals through ‘ten’).

TABLE 10.8 TOCHARIAN NUMBERS FROM ‘ONE’ TO ‘TEN’

	Toch. B nom.	Toch. B acc.	Toch. A nom.	Toch. A acc.	Toch. B gen.
‘one’ m.	še	šeme	<i>sas</i>	<i>šom</i>	(see below)
‘one’ f.	sana ~ somo	sanai ~ somo	<i>sām</i>	<i>šom</i>	
‘two’ m.	wi	wi	<i>wu</i>	<i>wu</i>	
‘two’ f.	wi	wi	<i>we</i>	<i>we</i>	
‘three’ m.	traï	traï	<i>tre</i>	<i>tre</i>	
‘three’ f.	tarya	tarya	<i>tri</i>	<i>tri</i>	täryāmts
‘four’ m.	štwer	štwer	<i>štwar</i>	<i>štwar</i>	
‘four’ f.	štwāra	štwāra	<i>štwar</i>	<i>štwar</i>	
‘five’	piś	piś	<i>pāñ</i>	<i>pāñ</i>	piśamts
‘six’	škas	škas	<i>šäk</i>	<i>šäk</i>	škässamts
‘seven’	šukt	šukt	<i>špät</i>	<i>špät</i>	
‘eight’	okt	okt	<i>okät</i>	<i>okät</i>	oktamts
‘nine’	ñu	ñu	<i>ñu</i>	<i>ñu</i>	
‘ten’	šak	šak-	<i>šäk</i>	<i>šäk</i>	

It is interesting to look closely at the paradigm for ‘one’ in both languages to see how the two languages have innovated with the material common to them so as to come up with quite different paradigms. The more usual forms of ‘one’ in the feminine singular, *sana/sanai/sām*, start from the PIE neuter singular \*sm̥ after final \*-m had become \*-n.

The non-nominative forms are generally from PIE \*somos (cf. English ‘same’). The palatalization of the word-initial \*s- presumably began in the masculine singular and plural, where it was analogically introduced for the same reason it was in the pronouns. In Tocharian A it has spread to the feminine as well. The rounding of PToch. \*-e- (< PIE \*-o-) before the bilabial -m- is regular in Tocharian A.

TABLE 10.9 THE DECLENSION OF ‘ONE’ IN THE TOCHARIAN LANGUAGES

	Toch. B m.	Toch. A m.	Toch. B f.	Toch. A f.
sg. nom.	še	<i>sas</i>	sanai ~ somo	<i>sām</i>
acc.	šeme	<i>šom</i>	sanai ~ somo	<i>šom</i>
gen.	šemepi	<i>šomāp</i>		
pl. nom.	šemi	<i>šome</i>	somona	<i>šomaṃ</i>
acc.	šemem	<i>šomes</i>	somona	<i>šomaṃ</i>
gen.	šememts	<i>šomešši</i>	somonaṃts	

The meaning of the plural of ‘one’ is ‘some’ (cf. the identical phenomenon in the Romance languages). Other numbers may also show a plural under the right conditions; A *šäptänt* ‘sevens’, i.e., ‘weeks’; or B *škanma* ‘tens’, i.e., ‘decades’; B *käntenna* ‘hundreds’.

In Tocharian B the second decade is composed of the numbers ‘one’ through ‘nine’ compounded with ‘ten’: *šak-še* ‘eleven’, *šak-wi* ‘twelve’, *šak-trai*, *šak-štwer*, *šak-pis*, *šak-škas*, *šak-šukt*, *šak-ñu*. The structure is essentially the same in Tocharian A, save for the addition of a conjunction -pi; thus *šak-wepi* ‘twelve’ (f.), *sak-nupi* ‘nineteen’, etc. The decade numbers, ‘twenty’ through ‘ninety’, show (partly opaque) derivatives of the numbers ‘two’ through ‘nine’: B *ikām* [A *wiki*] ‘twenty’, *täryāka* [*taryāk*] ‘thirty’, *štwārka* [*štwarāk*], *pišāka* [*pñāk*], *škaska* [*säksäk*], *šuktānka* [*šäptuk*], *oktānka* [*oktuk*], *ñumka* [*nmuk*]. As can easily be seen a lively analogy has been at work in both languages, but in quite different ways. The suffix reflects a PIE \*-(d)komt, which, at an early point in its history became pre-Tocharian \*-kom. The \*-om part was identical with the usual neuter singular noun suffix, whose singularness was felt to be inappropriate for these numbers and changed to the neuter plural \*-ā. Thus PToch. \*-kā and thence, regularly, Toch. B -ka, A -k. The word for ‘hundred’ is inherited: *kante* [*känt*] (< PIE \*kṛntóm), as is probably the word for ‘thousand’: *yaltse* [*wälts*], but higher numbers are borrowed: *tmāne* [*tmām*] ‘ten thousand’ and *kor* [*kor*] ‘ten million’.

The ordinal numbers from ‘second’ to ‘tenth’ are extremely transparent derivatives of the corresponding cardinal number and the suffix -te (Toch. A -t): *wate* [*wät*] ‘second’, *trite* [*trit*] ‘third’, *štarte* [*štärt*], *piñkte* [*pänt*], *škaste* [*škäst*], *šuktante* [*šäptänt*], *oktante* [*oktänt*], *ñunte* [\*ñunt], *škante* [*škänt*] ‘tenth’. In Tocharian B the process continues with *ikante*, but no higher ordinal numbers are attested. In Tocharian A we do not have ‘twentieth’, but there are *taryākiñci* ‘thirtieth’, *štwārākiñci* ‘fortieth’, and *säkskiñci* ‘sixtieth’ with the suffix -ñci. The word for ‘first’ is Toch. B *parwe/pärwešše* (cf. Sanskrit *pūrva-*), Toch. A *maltowinu* (but compare A *pärwat* ‘first-born son’).

VERBS: The Tocharian verb is the most complex part of Tocharian morphology. A given verb can distinguish two VOICES: active and medio-passive (intransitive,



reflexive, or passive); two ASPECTS: imperfective and perfective (but only in the past, in the distinction of imperfect [impf.] and preterit [pret.]); three MOODS: indicative (present and past), subjunctive (mostly with future meaning in main clauses, while signaling uncertainty or possibility in subordinate clauses), and imperative; NUMBER: singular, dual, and plural; and three PERSONS: first, second, and third. In these distinctions it is fully as complex as the “classical” Indo-European languages such as Sanskrit, Greek, or Latin. In addition, there is an INFINITIVE (made from the subjunctive stem in Tocharian B, from the present stem in Tocharian A); two VERBAL ADJECTIVES, one from the present stem, signaling necessity, and the other from the subjunctive stem, signaling possibility; a PRETERIT PARTICIPLE; and an AGENT NOUN. The verbal adjectives can be combined with ‘to be’ to create modal-like forms (‘one is to do X’, ‘one could do X’), and the preterit participle can also be combined with ‘to be’ to form a (rare) perfect (‘one has done X’). Most intransitive verbs, and many transitive ones, allow the derivation of a CAUSATIVE (“the window broke” > “John broke the window”, “John knows the rules” > “Tom showed/taught John the rules”). A small number of verbs, transitive or intransitive, have derived INTENSIVES, e.g., *kāln’ä/e-* ‘resound’ > *kālnāsk’ä/e-* ‘howl [of the wind]’, *oksó-* ‘awaken’ [intr.] > *āksāsk’ä/e-* ‘lie awake’.

The complexity of Tocharian verbal morphology is wholly consistent with that of its Proto-Indo-European ancestor. Not only is the system clearly Proto-Indo-European, but its expression is largely inherited as well. Among present stems, for instance, we find descendants of PIE *e/o*-presents, also *ye/o-*, *skel/o-*, *se/o-*, *eh<sub>2</sub>-*, *neh<sub>2</sub>-*, *nh<sub>2</sub>ye/o-*, and nasal-infix presents, as well as old denominatives reflecting PIE *\*-n-ye/o-* and *\*-eh<sub>2</sub>-ye/o-*. Thus, we have, using Tocharian B examples, *pārä/e-* (< *\*b<sup>h</sup>er-e/o-*) ‘bear’, *ceppi(ye)-* (< *\*topye/o-* [with innovative word-initial palatalization]) ‘step forth’, *pāsk’ä/e-* (< *\*peh<sub>2</sub>-ske/o-*) ‘guard’, *luks’ä/e-* (< *\*luk-se/o-*) ‘illuminate’, *lākā-* (< *\*l<sub>g</sub>-eh<sub>2</sub>-*) ‘see’, *tārknā-* (< *\*T<sub>ṛ</sub>K-ne-h<sub>2</sub>-*) ‘release’, *māntāññä/e-* (< *\*m<sub>ṃ</sub>t-ṅh<sub>2</sub>-ye/o-*) ‘harm’, *pīnkā-* (< *\*pi-n-k-*) ‘paint, write’, *kwipeññä/e-* (denominate to *kwipe* ‘shame’) (as if pre-Tocharian < *\*kwipen-ye/o-*) ‘be ashamed’, *klautko-* (denominative to *klautke*) (as if < *\*klowt<sup>k</sup>-eh<sub>2</sub>-ye/o-*) ‘change’. The Tocharian subjunctive shows many of the same formations, and, indeed, there are many instances where the present stem and subjunctive stems are identical. The most common subjunctive, however, is a formation with suffixal *-ā-* (of unclear and perhaps multiple origins). The present stems in the previous list are accompanied by the following subjunctives: *pārä/e-* (suppletive subj. *kāmā-*), *ceppi(ye)-* (subj. *tāppā-*, pret. *\*tāppā-*), *pāsk’ä/e-* (subj. *pāsk’ä/e-* [identical to present], pret. *pāššā-*), *luks’ä/e-* (subj. *luk’ä/e-*, pret. *lyāuk(sā)-* ~ *lāuksā-*), *lākā-* (subj. *lākā-*, pret. *lyākā-*), *tārknā-* (subj. *tārkā-* ~ *tārkā-*, pret. *cārkā-* ~ *tārkā-*), *māntāññä/e-* (subj. *māntā-* ~ *māntā-*, pret. *māntā-*), *pīnkā-* (subj. *pāikā-*, pret. *pāikā-*), *kwipeññä/e-* (subj. *kwipeññä/e-*, pret. *kwipeññā-*), *klautkó-* (subj. *klāutkā-*, pret. *klāutkā-*). Unlike the other non-Anatolian branches of Indo-European, Tocharian shows no traces of “long vowel” subjunctives, i.e., those subjunctives formed by adding the thematic vowel, in its role as a subjunctive marker, to the thematic vowel to give subjunctives in *\*-ē/ō-*. Nor are there thematic optatives in *\*-o-ī-* (< *\*-o-ih<sub>1</sub>-*). Rather, we have subjunctives and optatives without the thematic vowel, namely, subjunctives in *\*-e/o-* and optatives in *\*-ī-*, thus *pāsk’ä/e-*, both present and subjunctive < *\*peh<sub>2</sub>sk-e/o-* and optative *pāsk’i-* (< *\*peh<sub>2</sub>sk-ih<sub>1</sub>-*). The absence of “long vowel” subjunctives and thematic optatives in *\*-o-ī-* is clearly an archaic trait.

TABLE 10.10 CONSPECTUS OF THE TOCHARIAN VERB (Tocharian A forms in italics)

non-causative					
	<i>active</i>			<i>medio-passive</i>	
<i>tense</i>	NON-MODAL	MODAL		NON-MODAL	MODAL
present	kārsanaṃ [kārsnāš]	kārsam [subj.] [krasāš]		kārsanatār [kārsnātār*]	karsatār [subj.] [kārsātār]
non-present	kārsanoy [Impf] [šārs]	karsoy [Opt] [kārsi]		kārsanoytār [impf.] [kārsnitār*]	karsoytār [opt.] [kārsitār*]
preterit	šarsa [šārs]			kārsāte [kārsāt]	
preterit participle	kārsau [kārso]				
<i>non-finites</i>					
infinitive	[kārsnātsi]	karsatsi			
gerundive	kārsanalle [kārsnāl]	karsalle [kārsāl]			
abstract		karsalñe [kārsālune]			
<i>m</i> -participle	kārsanamane* [kārsnām*]				
<i>nt</i> -participle		kārsauca [kārsnānt]			
<i>causative</i>					
<i>tense</i>	NON-MODAL	MODAL		NON-MODAL	MODAL
present	šarsāššām [šārsāš]	šarsāššām [subj.] [šārsāš*]		šarsāštār* [šārsāštār*]	šarsāštār* [subj.] [šārsāštār*]
non-present	šarsāšši [impf.] [šārsši]	šarsāšši [opt.] [šārsāšiš*]		šarsāššitār* [impf.] [šārsšitār]	šarsāššitār* [opt.] [šārsāšitār*]
preterit	šārsa [šāšārs]			šārsate* [šāšārsāt*]	
preterit participle	šešārsu [šāšārsu]				
<i>non-finites</i>					
infinitive	[šārsāssi]	[šārsāstsi*]			
gerundive	šarsāššälle [šārsāšāl*]	šarsāššälle* [šārsāšāl*]			
abstract		šarsāššālñe* [šārsāšālune*]			
<i>m</i> -participle	šarsāškemane* [šārsāšmām]				
<i>nt</i> -participle	šarsāššeñca*				

NB: The finite forms are all third person singulars; the asterisk signifies a non-attested form/category (in all such cases, however, the reconstruction is certain). Here and elsewhere 'ā is an -ā-, which causes obligatory palatalization of a preceding palatalizable consonant. Thus, in this case, we get alternation between *pāššā-* and *pāške-* (e.g., third person singular *pāššām* and third person plural *pāškeṃ*, ultimately representing PIE \*peh<sub>2</sub>sketi and \*peh<sub>2</sub>konti respectively).

## SOME TOPICS OF WORD FORMATION

As in other Indo-European languages, Tocharian has a rich set of derivational processes. The commonest of these are the creation of denominative verbs, adnominal adjectives and possessive adjectives, abstract nouns, and nominal compounds.

**DENOMINATIVE VERBS:** There are two ways by which denominative verbs may be created, by suffixation or by compounding with the verb *yām-* [A *ypa-*] ‘do, make’. The process of suffixation is inherited but probably no longer productive in either language. With just enough examples to assure us of its existence, there was a process by which a (PIE) *o*-grade noun might take the verbal suffix *-ā-* to form a new verb. Thus, the Toch. B noun *klenke* ‘vehicle’ gave *klānkā-* ‘travel by vehicle’ or B *klautke* [A *lotāk*] ‘change’ gave B *klāutkā-* [A *lotkā-* (new analogical present *lotānkā-*)] ‘change’ (vb.). A bit more commonly, we have the suffix *-ññā/e-* [A *-iññā/a-*] (historically a denominative to *n*-stem nouns, i.e., PIE *\*-n-ye/o-*); thus, B *tan̄kw* [A *tuñk*] ‘love’ gives B *tānkwaññā/e-* [A *tuñkiññā/a-*] ‘love’ (vb.). Far, far more common than either of these “suffix solutions” is to combine a noun with the verb *yām-* [*ypa-*], e.g., B *apākārtse yām-* ‘be visible, be manifest’, B *saim yām-* [A *sem ypa-*] ‘take refuge’. Of the approximately 625 verbs known in Tocharian B, 54 are denominatives made with *yām-*; the numbers in Tocharian A would appear to yield a similar ratio.

**ADNOMINAL AND POSSESSIVE ADJECTIVES:** Adnominal adjectives are those that mean something like ‘pertaining to X’ or ‘consisting of X’, while possessive adjectives are those that mean something like ‘having X’. In each case the ‘X’ is a noun. Thus, B *kokale* [A *kukāl*] is ‘chariot/wagon’, and the adnominal B *kokaleṣṣe* is ‘pertaining to a chariot/wagon’, while the possessive adjective B *kokaletstse* is ‘having a chariot/wagon’ (nominalized as ‘charioteer, carter’) or B *stām* [A *štām*] ‘tree’, where the adnominal B *stamaṣṣe* [A *štāmaṣi*] is ‘pertaining to a tree’ and B *stanātstste* (built to the plural *stāna*) is ‘possessing trees’. Adnominal adjectives are essentially universally productive in both languages and play an interesting syntactic role (cf. *infra*); the possessive adjectives are common in Tocharian B but quite rare in Tocharian A.

**ABSTRACT NOUNS:** Both Tocharian languages have fully productive means by which to create abstract nouns from adjectives and nouns. In Tocharian B there is *-āññe* (less commonly *-uññe* or *-auñe*), and in Tocharian A there is *-une* (less commonly *-one*). (The Tocharian B and A suffixes are presumably ultimately related, but the relationship is not well explained.) Thus, in Tocharian B we have *astare* ‘pure’ > *astarñe* ‘purity’ [A *āštār* > *āštrone*], B *parkār* ‘long’ > *pārkarñe* ‘length’ [A *pärkär* > *pärkrone* ‘length’], or B *lānt* ‘king’ (acc.) > *lantuñne* ‘kingship’ [A *lānt* > *lāntune*]. The creation of abstract nouns from adjectives includes creation from the productive deverbal adjectives in *-lle* [A *-l*]. Usually it is the verbal adjective built on the subjunctive stem (potential gerundive) that forms the basis of the verbal abstract noun, but sometimes one also finds abstracts built to the verbal adjective derived from the present stem (the gerundive of necessity): B *tsrālñe* ‘separation’ based on the subjunctive stem *tsārā-* but also *tsrelñe* ‘id.’ based on the present stem *tsäre-* or A *kālpālune* (from the subjunctive) and *kālpnālune* (from the present), both ‘achievement’.

**COMPOUNDS:** The Tocharian languages are rich in nominal compounds. Essentially all types of compounds that are familiar in other Indo-European languages can be found. There are copulative compounds (the *dvandvas* of Sanskrit grammatical tradition), e.g., B *ñem-kālywe* ‘fame’ [A *ñom-klyu*] (‘name-fame’), B *pācer-mācer* [A *pācar-mācar*] ‘parents’ (‘father-mother’), B *śwātsi-yoktsi* [A (pl.) *śwātsi-yoktsintu*] ‘food and drink’. There are also numerous examples of determinative compounds acting as adjectives (the

Sanskrit grammarians' bahuvrīhis): as B *astre-pālsko* 'with pure spirit', B *tāryā-yākne* 'three-fold', B *totkā-yārm* 'of small measure' [cf. A *tsru-yārm* 'id.'], B *treya-meskeṃ* 'having three joints' [cf. A *tri-asamkheṣi* 'pertaining to the three ages'], B *pärkre-šaul* 'long-lived', B *ška-maiyya* 'having ten powers' [cf. A *ška-tempeyum* 'having ten powers'], B *šeme-yārm* 'having a single measure', B *še-šuke* 'having a single taste', A *šoma-pācar* 'having the same father'. Sometimes the bahuvrīhis are provided with an adjectival derivational suffix (B *-tstse*, A *-(u)m*) as well, e.g., B *orotstse-cämpamñetstse* 'having great capabilities', B *yolo-pilkotstste* 'with bad insight', A *ñom-kälywäts* 'famous', A *ñom-klyum* 'famous', A *ška-tempeyum* 'having ten powers'. Finally, we find substantial numbers of endocentric determinative compounds (tatpuruṣas), though this latter sort of compound may be less numerous than, say, in Germanic, since they have heavy competition in the form of adnominal adjectives plus nouns (cf. supra and infra). With a regular verbal derivative (participle, verbal noun, agent noun) as the second member we have, for instance, B *osta-šmeñca* 'householder', B *kärtse-akšu* 'well instructed', B *kärtse-rita* 'seeking good', B *telki-yāmor* 'sacrifice'. With nouns not directly derived from verbs as second members we have B *kärtse-palsko* [A *kāswa-pāltsāk*] 'having a good thought', B *orotstse-pācer* 'grandfather', B *kašār-wātsi* 'kāšāya-clothing' [cf. A *kāsāri wsāl* 'id.'], B *kuñi-mot* 'grape-alcohol' (i.e., 'wine'), B *kaum-ñakte* [A *kom-ñkät*] 'sun-god', B *pañakte ~ pūd-ñakte* [A *ptā-ñkät*] 'Buddha', B *poysi* 'Buddha' (i.e., 'all-knowing') [cf. A *puk-knānmām* 'id.'], B *šañ-añm* [A *šñi-āñcām*] 'oneself', B *šañ-šaumo* 'relative, retainer, follower'.

## SYNTAX

The Tocharian languages are typical Indo-European languages in that they have a relatively "free" word order because the rich morphology will often tell the relationship of a noun, say, to other nouns or to the verb in a sentence. The result of this morphological robustness is that rarely is it the case that, unlike the situation in English, some particular constituent of a sentence (e.g., subject, verb, or adjective) has to be in a particular syntactic position in a sentence. Still, the order of elements in a Tocharian B sentence is not altogether random. It is true, far more often than not, that in independent clauses the semantically neutral, or unmarked, word order in Tocharian prose is subject-object-verb. A Tocharian B example, from a letter with no literary pretensions, is *Šilarikite* [sic] *Āryawarmem tsamo yšuwarsa prekšām* 'Šīlarkšita inquires with great friendship of Āryavarman'. In copular (or "nominal" sentences) the order is subject-predicate-(BE), where the BE, at least in the present tense, is often not present (e.g., from another letter, *sānkantse ayāto nesaññe ste* 'the situation of the community is satisfactory' with BE present, and, in a literary text, [c]m(e)laššai ytāri yātsi tumem (pre)[k]e : 'therefore [it is] time to go the way of births' with BE absent).

Particularly in poetry different orders appear *metri causa* or for literary effect. But even in non-literary texts, emphasis or focus or new information may be shown by fronting the syntactic constituent involved. It is almost certainly the case that *any* element of a clause may be fronted. Because of both emphasis and focus criteria, imperative verbs are commonly, though not universally, fronted, e.g., *pluwas wesi šak meñatse-ne wai parāce komne ploriyaceṃ* 'send us, on the tenth of the month and the day of recitation, musicians'. Also normally fronted are interrogative and relative pronouns. With interrogative pronouns: *k<sub>i</sub>ce klautkesa aišalyi* 'by what behavior [are they] to be recognized?', • *k<sub>i</sub>cesa akalkāntats yatalyñe mäsketrä* : 'by which is the taming of desires?'; with relative

pronouns: *ket ait yoktsi sām̄tk=onwaññe pelaiykneṣṣe lyaitkem teki mantanta šāp srukentār cai* • ‘to whom thou givest the righteous, immortal medicine to drink, they will avoid sickness and never die’, *k<sub>u</sub>ce te mant wñāwa, tu ñke weñau anaiśai* ‘what I have so said, that will I now speak clearly’.

Constituents can be shifted rightward as well. Right-detachment is a process whereby appositional phrases or other adjuncts are moved to the end of a grammatical sentence, e.g., *ysāre lac piś cakanma wi tauwā šwer; wāsokās āka laś cāk piś tom* ‘Wheat went out, five *cāks*, two *tom*; for the [making of] *wāsok* [= beer?] barley went out, one *cāk*, five *tom*’. In both clauses we have intransitive sentences (‘wheat went out’, ‘barley for beer went out’) where the subject has been expanded by the addition of measurements, but those measurements come after the verb.

The usual place for an adjective or genitive is before the noun it modifies. Non-literary monastic records provide abundant examples, e.g., *olyīškāṣṣe Śinkeñe werwiyesa* ‘for Śinke’s *olyiška*-garden’, *wace kraī* ‘in the second market/on the second market-day’, *ptamaṣṣe werwiyesa* ‘for the stupa’s garden’ (for the adjective), or *stere Nānatewetse ṣotri* ‘the sign of the elder Jñānadeva’, *kapyāres śwasiṣṣe* ‘food-provisions for the monastic workers’ (for the genitive). Other kinds of modifiers, e.g., prepositional phrases, also tend to proceed, e.g., *yṣiñe cokiś ṣalywe* ‘oil for the night lamp’. But even in documents such as these we find the reverse order, e.g., *ceynas pito koromñe* ‘the price of their [transportation] by mules’. In more literary documents, the orders N–A and N–G occur freely. Particularly in more literary texts the modifier and noun may be separated from one another (“distracted” element underlined), e.g., *se wcepī āke ślokantse* ‘the end of the second stanza’, *wesāñ ñake ṣarnene kekamu nest* ‘into our hands thou hast now come’, *snai keś lkāṣṣām kektseñān* ‘he sees countless bodies’. A rather extreme case of distraction, and in general rearrangement of expected word order, is *keklyauṣormem krent ñem-kālywe wroccu wlo tañ* ‘having heard of thy good renown, [O] great king’, where *tañ* belongs with *krent ñem-kālywe* and \**tañ krent ñem-kālywe* would normally precede *keklyauṣormem*. Nevertheless, distraction is possible even in a non-literary letter, e.g., *sāñkantse ayāto nesaiññe ste* ‘the situation of the community is satisfactory’.

There are three kinds of dependent clauses in Tocharian B: (1) those that modify a particular noun or pronoun, (2) those that function as part of the verb phrase as an adverb of some sort, and (3) those that function as a subject, direct object, or predicative nominative. The first group is composed of relative clauses; the second of temporal, local, manner, causal, and result clauses; and the third group of noun clauses. Dependent clauses almost invariably precede the main clause. Relative clauses are a partial exception in that we find both orders without apparent difference in meaning; some follow (relative pronoun bolded): *cauwak yakne énkaskemttār {mūkcau procer énsate}* ‘we take that very way that [our] brother has taken’, */// pepār̄kormem yamor krentā {mūkcpī} okonta wārpānoytrā* ‘having asked after the good deed whose fruits he enjoyed’, but some precede: *{k<sub>u</sub>ce cai makci enkoṣṣem} pilkonta* ‘these insights which they themselves [have had]’. Most commonly, however, relative clauses are “correlative,” e.g., *{k<sub>u</sub>ce te mant wñāwa}, tu ñke weñau anaiśai* ‘what I have so said, that will I now speak clearly’.

Dependent clauses are marked by a clause-initial subordinator (relative clauses are again an exception in that the subordinator, the relative pronoun, is not always clause initial). The main clause, too, may be marked by a clause-initial particle, e.g., *{ente palsko tsāñkam-ne} ot śwātsi yoktsi kāl̄pāsām* ‘whenever the spirit arises in him, then he finds

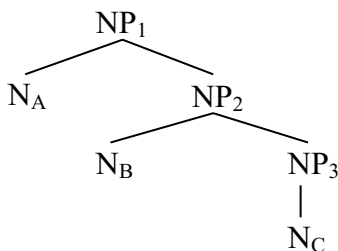
food and drink'. Comparative clauses and correlative relative clauses are (almost) always so marked; other types show a main clause particle 20% or 30% of the time.

As in other Indo-European languages with morphological verbal moods, the Tocharian subjunctive is commonly found in dependent clauses. Where both subjunctive and indicative can be found, the subjunctive reflects an indefinite or unspecified subject/condition or is future-oriented, or both. The optative in these dependent clauses is used either as the past equivalent of the subjunctive (by a Tocharian version of the sequence of tenses rule) or when the dependent clause is offered as unreal, e.g., in contrary-to-fact conditional clauses ("if I were rich . . .") or in relative clauses where the antecedent is denied ("there is no one who . . .", e.g., B *mā sū nesām {k<sub>u</sub>se onwaññe tākoy}* 'there is no one who is [lit. may/might be] immortal').

One of the characteristic syntactic phenomena of Tocharian is the widespread presence of adnominal adjectives. To a much greater degree than is the case for most Indo-European languages, Tocharian has a productive grammatical rule whereby a modifying genitive noun may be converted into an adnominal adjective that, then, will agree in gender, number, and case with its head noun. Thus, one can say either B *slentse* [GEN.] *trōnkne lyam=ompalskoññe* 'he sat [pret.] in meditation in a hollow of the mountain' or *sleye* [lege *šliye*] [ADJ.] *gune cau šamy ompalskoññe* 'he sat [impf.] in meditation in that mountain cave', *rūpn=énwentse* [GEN.] 'in the shape of a man', but *secakāññe* [ADJ.] *rupsa* 'in the shape of a lion', or further B *skrenantse* [GEN.] *paruwa* 'crow-feathers', but *skrense* [ADJ.] *paiyye* 'crow-foot'. These adnominal adjectives may take one of three possible shapes: (1) *-iye* (inherited directly from PIE \*-iyo- and restricted to long-established inherited nouns, e.g., *ypīye* 'barley-' [from *yap* 'barley'], *wāstarye* 'camel-' [underlying noun unattested], *kewīye* 'bovine' (as a noun, 'butter') [*ke<sub>u</sub>* 'cow'], *patarye* 'paternal' [*pācer* 'father'], etc.); (2) *-ññe*, common only in Tocharian B, from PIE \*-n-yo- and largely restricted to animate nouns (e.g., *yākweññe* 'equine' [*yakwe* 'horse']); and (3) *-šše*, largely restricted to inanimates and dual or plural nouns of all degrees of animacy (e.g., *klautsaišše* 'pertaining to the ear', *klautsanešše* [A *klošnāši*] 'pertaining to the two ears'). However, the distinction is not rigid, and there are nouns that give rise to more than one denominal adjective, e.g., *yākweññe* and *yākwešše* [A *yukaši*], both 'pertaining to a horse'.

Unlike the situation in other Indo-European languages in which this formation exists, where it can be used only when the modifier is singular and itself otherwise unmodified, Tocharian allows the modifying noun to be of any number (as also in Luvian among the Anatolian languages), and it, in turn, may be modified (thus, we have the two adjectives *klautsaišše* 'of the ear' and *klautsanešše* 'of the two ears', *lāklešše* 'pertaining to suffering' (from *lakle* 'suffering') and *lāklentašše* 'pertaining to sufferings' [Toch. A *klop/klopant* 'suffering(s)', whence both *klopaši* and *klopāntwāši*]). In many cases the genitive and adnominal adjective are semantically identical, or substantially so. In other cases the use of the genitive may signal a specified noun, while the adjective may be more generic, e.g., *lānte warqšāl* 'the king's power', but *lantuññe warqšāl* 'royal power'. Animacy is also a factor: the more animate the modifying noun, the more likely that it will remain in the genitive and vice versa.

Even more astonishing is the behavior of this phenomenon when it occurs in noun phrases with more than one level of modifier-embedding. A noun phrase (NP<sub>1</sub>) can consist of a noun (N<sub>A</sub>) and a modifying noun phrase (NP<sub>2</sub>), which, in turn, can consist of a noun (N<sub>B</sub>) and noun phrase (NP<sub>3</sub>), and so on. In the diagram below, either or both of N<sub>B</sub> and N<sub>C</sub> can be a noun in the genitive or a derived adnominal adjective.



If  $N_C$  is a derived adnominal adjective, the adjective agrees in gender with the noun  $N_B$ , even if the noun has itself been changed into an adnominal adjective (agreeing with the head noun,  $N_A$ ). Let's look at an example, this one from a list of medicinal recipes: *läksañai* [ADJ., F. ACC. SG.] *klautsaişşe* [ADJ., M. NOM. SG.] *şpel* [NOUN, M. NOM. SG.] 'poultice of fish ears/gills'. The  $N_B$  slot is occupied by the adjective *klautsaişşe*, which is regularly derived from the feminine noun *klautso* 'ear/gill'. It is in the masculine nominative singular form because it agrees with  $N_A$ , *şpel* 'poultice', a masculine nominative singular noun. The  $N_C$  slot is also occupied by an adjective, *läksañai*, which is regularly derived from *läks* 'fish'. However, it is in the feminine singular genitive/accusative form because it agrees with the feminine genitive singular of *klautso*, namely, *klautsaintse*, which underlies the surface-structure *klautsaişşe*. Particularly interesting is the double example such as *mäkce<sub>u</sub> ykeşşa kektseñe tã, kenaşşe satãşlñe* 'whatever place the body [has], exhalation [has] that [place on] earth'. Here we have *mäkce<sub>u</sub>*, which is masculine, agreeing with the underlying *ike* 'place' (m.), even though *ike* has become *ykeşşa*, a feminine adjective, to agree with the feminine noun *kektseñe* 'body'. Conversely, *tã<sub>u</sub>* is feminine to agree with the feminine noun *keñ* 'earth', which has been replaced by the adjective *kenaşşe* (masculine to agree with *satãşlñe* 'exhalation').

## FURTHER READING

Supporting the notion that "Tocharian" is a correct name for those languages so-called is Adams 2000. It is a review of the data and analysis that probably justify grouping the classical *Tokharoi*, the Uyghurs' *Tuyrï*, and the *Tocharians* into a single historical "ethnic complex." It is certain that not all investigators will agree at every point. Balanced reviews of the prehistory of the Tarim Basin are Barber's (1999) *The Mummies of Ürümqi* and J. P. Mallory and Victor M. Mair's (2000) *The Tarim Mummies*. Both books are readable and very informative accounts of the prehistory of the Tarim Basin, its peopling, and its culture.

Overwhelmingly the best place to go for the Tocharian texts themselves is CEToM 2011–, *A Comprehensive Edition of Tocharian Manuscripts*. The database is as yet incomplete but is designed to give digital access to very nearly all Tocharian manuscripts, accompanied, where possible, by (tentative) transcriptions and commentary; this will be a resource for Tocharianists that is unlike that of any other Indo-European group. Accessible now are the texts housed in London, those housed in Berlin, and many of those housed in Paris.

Krause and Thomas 1960–1964 is a two-volume work containing a serviceable if brief grammatical sketch (phonology and morphology) of both languages, along with selected texts and glossaries from both. Sieg et al. 1931 is an early work that provides what is still in some respects the best and fullest grammar of the two languages, but clearly focused on Tocharian A. It is dated but still very serviceable, particularly with regard to morphology.

Like Krause and Thomas 1960–1964, Pinault’s chrestomathy (2008) contains readings and grammatical analysis, along with considerable discussion of the development of Tocharian from Proto-Indo-European. Peyrot (2008) provides the primary discussion and analysis of the internal, half-millennium-long history of Tocharian B. Malzahn 2010 and Peyrot 2013 together form an exceptionally thorough morphological description of the Tocharian verb. Thomas 1957 is a thorough examination of Tocharian aspect. None of these is for the neophyte. A full-fledged syntax of either or both languages is clearly a desideratum. (But see now Adams 2015.)

The Tocharian A glossary of Krause and Thomas (1960–64) remains the best overall dictionary for that language, though it is supplemented by the incomplete dictionary of Carling et al. (2009). Adams 2013 is the most complete dictionary, synchronic or diachronic (it is both), of either Tocharian language, though by no means the “final word.”

## REFERENCES

- Adams, Douglas Q. 2000, “Some observations of peoples, places, and languages in the Taim Basin in the first millennium AD,” *Tocharian and Indo-European Studies* 9: 1–28
- Adams, Douglas Q. 2013, *A Dictionary of Tocharian B*, 2nd revised and enlarged edition, Rodopi, Amsterdam–Atlanta
- Adams, Douglas Q. 2015. *Tocharian B, A Grammar of Syntax and Word-Formation*. Institut für Sprachen und Literaturen der Universität Innsbruck, Innsbruck.
- Barber, Elizabeth Wayland 1999, *The Mummies of Ürümqi*, W. W. Norton, New York–London
- Carling, Gerd; Pinault, Georges-Jean & Winter, Werner 2009, *Dictionary and Thesaurus of Tocharian A, Volume I: A–J*, Harrassowitz, Wiesbaden
- CEToM (2011–), *A Comprehensive Edition of Tocharian Manuscripts* [www.univie.ac.at/tocharian/Krause](http://www.univie.ac.at/tocharian/Krause),
- Krause, Wolfgang & Thomas, Werner 1960–64, *Tocharische Elementarbuch, Bände I, II*, Carl Winter, Heidelberg
- Mallory, James P. & Adams, Douglas Q. (eds.) 1997, *Encyclopedia of Indo-European Culture*, Fitzroy Dearborn, London–Chicago
- Mallory, J.P. & Mair, Victor M. 2000, *The Tarim Mummies*, Thames and Hudson, London
- Malzahn, Melanie 2010, *The Tocharian Verbal System*, Brill, Leiden
- Matzinger, Joachim 2006, *Der Albanische Text “Mbsuame e Krështerë” (Dottrina Cristiana) des Lekë Matrënga von 1592: Eine Einführung in die albanische Sprachwissenschaft*, (Jenaer Indogermanische Textbearbeitung, 3.), J. H. Röhl, Dettelbach
- Peyrot, Michaël 2008, *Variation and Change in Tocharian B*, (Leiden Studies in Indo-European 15.), Rodopi, Amsterdam–New York
- Peyrot, Michaël 2013, *The Tocharian Subjunctive: A Study in Syntax and Verbal Stem Formation*, Brill, Leiden
- Pinault, Georges-Jean 2008, *Chrestomathie tocharienne: Textes et Grammaire*, Peeters, Leuven–Paris
- Sieg, Emil; Siegling, Wilhelm & Schulze, Wilhelm 1931, *Tocharische Grammatik*, Preussische Akademie der Wissenschaften, Göttingen
- Thomas, Werner 1957, *Der Gebrauch der Vergangenheitstempora im Tocharisch*, Harrassowitz, Wiesbaden
- Winter, Werner 1955, “A linguistic classification of tocharian B texts,” *Journal of the American Oriental Society* 75: 216–225
- Winter, Werner 1962, “Nominal and pronominal dual in tocharian,” *Language* 38: 111–134





Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

PART 11

# **BALTO-SLAVIC**



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

# BALTO-SLAVIC

*Steven Young*

The similarities in structure and lexicon, some quite striking, between the Baltic and Slavic branches of the Indo-European language family have long drawn scholarly attention. But the precise nature of the relationship between the two branches has been a source of debate for nearly as long. The prevailing view up to the early nineteenth century was that the Baltic languages were a kind of Slavic. It was through the work of the comparativists Franz Bopp and August Pott that Baltic was finally recognized as a language family in its own right (Petit 2004: 17–19). The classic presentation of Balto-Slavic as an intermediate branch of the Indo-European language family, with Baltic and Slavic as parallel sub-branches, is reflected in August Schleicher’s *Stammbaum*, or family-tree, approach, in which “slawolitausch” (Balto-Slavic) develops from an earlier “slawodeutsch” (Petit 2004: 19–20). Comparative evidence for a Balto-Slavic grouping was later supplied by Karl Brugmann (1904: 18) in the form of eight shared innovations in phonology and morphology, among them the development of syllabic sonorants such as *-iR-*, the use of pronominal *\*-yo-* to form definite adjectives, and the replacement of the inherited genitive singular form of *o*-stem nouns by the ablative.

The first to offer a principled challenge to the neogrammarian tenet of a post-Indo-European period of Balto-Slavic common development was Antoine Meillet in 1908 (cf. 1967: 59–67).<sup>1</sup> In Meillet’s view, the various agreements between the two language families reflect either inherited Indo-European archaisms (with both Baltic and Slavic developing from nearly identical Indo-European dialects) or later parallel developments in each of the branches, rather than shared innovations that would demonstrate a stage of linguistic unity.

Jan Endzelin’s 1911 study of the agreements and differences between Baltic and Slavic represents an attempt to reconcile Meillet’s approach with that of the neogrammarians. While Endzelin agrees with a number of Brugmann’s shared innovations, adding to these syntactic comparisons (predicate instrumental and genitive direct object in a negated sentence) and a substantial number of exclusive lexical agreements, he accounts for these by positing a period of prolonged language contact between neighboring Baltic and Slavic communities, leading to a degree of linguistic convergence, particularly in the lexicon. According to Endzelin, while we can justifiably speak of a common Balto-Slavic period (“эпоха совместной жизни”: 202), this does not imply a Balto-Slavic protolanguage, reflecting a stage of complete linguistic identity.

Additional arguments, both in support of and against a period of Balto-Slavic linguistic unity, were put forward over the course of the twentieth century. Among the notable attempts in support of a “Sprachgemeinschaft” is Reinhold Trautmann’s 1923 *Baltisch-Slavisches Wörterbuch*, in which Baltic and Slavic lexical correspondences are systematized. A shortcoming of this work is the frequent inclusion of external cognates, and etyma that are represented in either Baltic or Slavic, but not in both (Petit 2004: 26).

A review of the accumulated evidence in support of the notion of Balto-Slavic can be found in Stang 1966 (17–21). Shared phonological innovations include the development of acute and circumflex tone on non-short syllables (*ī̄*, *Vi/u*, *VR*): Lith. *kárvė* ‘cow’ : SCr. *kráva* (reflex of acute tone), Lith. *saũsas* ‘dry’ : SCr. *sûh* (reflex of circumflex tone); and the development of the syllabic sonorants *\*r̥*, *\*l̥*, *\*ŋ̥*, *\*ŋ̥* into the sequences *ir*, *il*, *im*,

*in*: Lith. *viršūs* ‘top’ : ORuss. *вѣрхъ*, *vilkas* ‘wolf’ : OCS *vlъkъ* (\*-ьl-), Lith. *dešimtas* ‘tenth’ : OCS *desętbъ* (\*-ьn-), OPruss. *insuwis* ‘tongue’ : OCS *jęзыкъ* (\*-ьn-) and (more controversially) after labiovelar stops into *ur*, *ul*, *um*, *un* (Vaillant 1950: 171): CS \*gьrdlo ‘throat’, Lith. *gurklỹs* ‘crop’ reflect a reduced grade \*gur- to the \*g<sup>w</sup>erh<sub>3</sub>- of CS \*žerti (OCS *požřęti* ‘to swallow, devour’), Lith. *gęrti* ‘to drink’. Compare also the set OCS *gьnati* (reduced grade) ‘to chase’ : *ženę* (full grade) ‘(I) chase’ and OPruss. *guntwei* ‘to chase, drive’, *gunnimai* ‘(we) chase’ and Lith. *genũ*, Latv. *džęnu* ‘(I) drive (cattle)’.<sup>2</sup>

To Stang’s examples of Balto-Slavic phonological innovations we can add the change of *e* to *o* before *u* when followed by a back vowel: Lith. *tãvas* ‘your’, Latv. *tavs*, compare Gr. *τεός*; OCS *novъ* ‘new’, compare Gr. *νέος* (Endzelin 1911: 84); the change of *-eu-* to *-iau-* (*-au-*) before a consonant:<sup>3</sup> Lith. *liãudis* ‘people, nation’, Latv. *lãudis* (nom. pl.) ‘people, folk’ : OCS *ljudьje* ‘people’ (\*lewd<sup>b</sup>-); Lith. *liaupsẽ* (\*lewb<sup>b</sup>-) ‘praise’ : OCS *ljuby* ‘love’.

Innovations in nominal derivation (see Stang 1966: 18) include deverbal abstract nouns in *-ima-*: Lith. *piešimas* ‘drawing’ : CS \*письмо ‘writing’; abstract formations in \*-ĩ-b-: Lith. *draugybė* ‘friendship’, Latv. *draudzība* : CS \*дружба; agent nouns in *-ã-jo-*, *-ẽ-jo-*: Lith. *artõjus* ‘plowman’, OPruss. *artoys* : OCS *ratajъ*; Latv. *šuvēja* (\*šjuu-ẽĩ-) with an exact match in Russ. *увей* ‘tailor’ (Endzelin 1952: 37); agent nouns in *-ik-o-*: Lith. *siuvinikas*, OPruss. *schuwikis* ‘shoemaker’ : ORuss. *šьвѣсь* ‘tailor, shoemaker’; the agentive suffix represented by dialectal and OLith. *-inỹkas* (standard *-iniņkas*), CS \*-ьникъ: Lith. dial. *arklinỹkas* ‘stableman, groom’ : OCS *dlъžьnikъ* ‘debtor’; and the diminutive suffix \*-ĩt(i)õ-: Lith. *sũnỹtis* ‘son (dim.)’ : OCS *otročĩstъ* ‘boy’.

This list can be supplemented by denominal adjectives in *-in-*: Lith. *krũvinas* ‘bloody’ : OSC *krъvньnъ* (Endzelin 1952: 38); diminutives in *-uk-o-*: Latv. *dļļuks* ‘son (dim.)’, OCS *synъkъ*; masculine derivations in *-in-* such as Lith. *ãvinas* ‘ram’ (compare *avis* ‘sheep’), Latv. dial. *avins*, OPruss. *awins* : OCS *ovьnъ*; and oblique cases of the first-person singular personal pronoun based on a stem \*me/on-, generalized from the genitive stem: Lith. gen. *manęs*, dat. *mãn*, acc. *manẽ*; Latv. gen. *manis*, dat. *man*, acc. *mani*; OPruss. dat. *menei* : OCS gen. *mene*, ORuss. dat./loc. *mъnẽ*; the reduced grade of the latter is matched by the dialectal (High) Latvian and Žemaitic Lithuanian dative *mun*.

Innovations in the inflectional and derivational morphology of the verb (Stang 1966: 18–19) include the *-ã* formant marking a preterit/aorist stem, often accompanied by the reduced grade of the root: Lith. *pirko* (pres. *peřka*) ‘bought’ : OCS *žьda* (pres. *židetъ*) ‘waited, expected’; the Baltic preterit in \*-ẽ: Lith. *vẽdẽ* ‘led’, compared with the Slavic imperfect of the type *veděaxъ* ‘(I) led’; the verbal suffix Lith. *-ãujV-* compared with Slavic \*-ũjV-, preterit Lith. *-ãvo* and Slavic *-ova*; and the distinctive formation of the present tense stem of ‘give’: OLith. *duosti* ‘gives’ : OCS *dastъ* (\*dõ-d(ã)-ti).

In addition, Baltic and Slavic share a large number of apparently exclusive lexical correspondences, which are not concentrated in any single semantic field. Wojciech Smoczyński (2001: 16–17), following Endzelin 1952 (44), finds at least two hundred, among them: Lith. *rankã* ‘hand’, Latv. *rũoka* : OCS *rъka*;<sup>4</sup> Lith. *rãgas* ‘horn’, Latv. *rags* : OCS *rogъ*; Lith. *liẽpa* ‘linden’, Latv. *liẽpa* : OCS *lipa*; Lith. *lẽdas* ‘ice’ : OCS *ledъ*; Lith. *parpartỹs* ‘fern’ : CS \*paportъ. These can be supplemented by the following comparisons from Endzelin 1952 (42ff.): Lith. *piršys* (nom. pl.) ‘chest (of a horse)’ : OCS *prъsi* (\*-ьr-) ‘chest, bosom’; Lith. *sietas* ‘sieve’, Latv. *siẽts* : SCr. *si̇to*; Lith. *mielas* ‘nice, sweet, dear’, Latv. *mĩļs* ‘sweet, dear’ : CS \*mĩľъ; Lith. *slãnka* ‘woodcock’, Latv. *slũoka* (with anomalous tone) : Russ. *слыка*; and the precise correspondence sets Lith. *vãrna* ‘crow’, Latv. *vãrna*, OPruss. *warne* : CS \*võrna (whatever the interpretation of acute in this form may be) and Lith. *vařnas* ‘raven’, OPruss. *warnis* : CS \*võrnъ.

Finally, there are a number of prepositions and preverbs with close or identical correspondences in the two language families: Lith. *už* ‘at, within, instead of, in return for’, Latv. *uz* ‘on, to’ : CS \*vъz- ‘in return for, (preverb) up, back’; Lith. *priẽ* ‘at, with, to’, OPruss. *prei* ‘at,

with, to' : CS \*pri 'at, with, by'; Lith. *nuõ*, Latv. *nùo* 'from', OPruss. *no, na* 'onto, against, over' : CS \*na 'on(to), in(to)'; Lith. dial. and OLith. *iž* (standard *iš*), Latv. *iz* 'out of' : CS \*jъz 'from, out of', with unexplained but matching initial \*i-. A refinement of Balto-Slavic lexical comparisons, drawing on progress in dialectal data, can be found in Otkupščikov 1988.

Perhaps the most compelling argument for a Balto-Slavic protolanguage, demonstrated on the basis of shared innovations, is to be found in the highly complex prosodic structures of both language families,<sup>5</sup> which typically agree in accentual paradigms and lexical membership in these paradigms, stress placement (including the reflexes of Hirt's Law), and phenomena associated with the rise of syllable tone (including the reflexes of Winter's Law).

The classic statement of tonal correspondences between Baltic and Slavic is reflected in Stang 1966 (125). In this traditional view, the correspondence set conventionally labeled acute, continuing Indo-European length (including length from loss of a laryngeal), shows a rising pitch in Slavic, Old Prussian and (together with "broken" tone) Latvian, but a falling pitch in (standard) Lithuanian. On the other hand, circumflex tone has a falling pitch in the former, but a rising pitch in Lithuanian. Stang (1966: 125) assumes a shift in tonal contours ("eine Umlegung von Tonverlauf") for Lithuanian. An attempt at resolving this anomaly is found in the work of Frederik Kortlandt (for example, 1985: 122ff.), who has argued that the various tonal contours are later developments in the individual languages, and that the broken-tone (glottalic) acute of Latvian and the Žemaitic dialects of Lithuanian is in fact the prosodic continuation of an earlier laryngeal. (Baltic) circumflex tone is simply the continuation of a non-laryngeal length (lengthened grade or vowel contraction).

An apparent lack of tonal correspondences between Baltic and Slavic can be found in stems belonging to the mobile accent paradigm (see below). Here the opposition of acute and (Balto-Slavic) circumflex was neutralized in Slavic; all such roots show "Slavic circumflex," which continues in Serbo-Croatian as long falling tone: SCr. *gláva* 'head', acc. sg. *glávu* ≠ Lith. *galvā, gálvą*, Latv. *galva*, with acute; SCr. *nāg, nāga* 'naked' ≠ Lith. *núogas, Latv. nuógs*, with acute; SCr. *žív, žíva* 'alive' ≠ Lith. *gývas, Latv. dzívs*, with acute. This development, first described by Antoine Meillet (1902), is known as Meillet's Law.

The set of Balto-Slavic acute stems of laryngeal origin is supplemented by the reflexes of Winter's Law, which likewise serves as an important argument for a Balto-Slavic linguistic unity (Derksen 2004: 82). Werner Winter (1976) showed that there is a relationship between instances of Balto-Slavic lengthened vowels and the distribution of stem-final Indo-European aspirated and unaspirated voiced stops; an original short vowel followed by an aspirated stop remains short in Balto-Slavic, while an original short vowel followed by an Indo-European voiced stop shows a lengthened (acute) vowel:<sup>6</sup> BSl. \*med- 'honey' < \*med<sup>h</sup>-, but \*ēd- 'to eat' < \*ed-. Winter's Law accounts for a number of cases that Stang had termed "theoretical metatony" (1970: 216), where the acute, rather than the expected circumflex (or short), occurs in a non-laryngeal base. The most comprehensive list of Balto-Slavic data supporting Winter's Law is undoubtedly Dybo 2002.

Baltic and Slavic also agree in showing unexpected root stress in a number of bases that are oxytonic in Greek and Sanskrit: Lith. *dūmai* 'smoke', Latv. *dūmi*, Slav. \*dŭmъ a.p. (a), but Skr. *dhūmā-*, Gr. θῆμός; Latv. *grīva* 'mouth of a river', CS \*gríva a.p. (a) 'mane', but Skr. *grīvā* 'neck'; Lith. *pilnas* 'full', Latv. *pīlns*, CS \*pŭlnъ a.p. (a), but Skr. *pūrṇā-*; Latv. *diēveris* (\*deh<sub>2</sub>iwer-) 'husband's brother', SCr. *děvēr*, but Skr. (RV) *dēvā* '(younger) brother of husband', Gr. δᾶήρ [\*δαίήρ] 'husband's brother'. This set of correspondences was first noted by Hermann Hirt (1892: 39), who characterized it (originally for Lithuanian) as "eine Akzentverschiebung bei den Worten mit gestossenem Vokal in der Stammsilbe . . ." 'a stress shift in words with a broken (glottalic) vowel in the stem syllable'. This statement, known as Hirt's Law, was revised in connection with laryngeal theory by V. M. Illič-Svityč (1963: 78–81); in Derksen's updated formulation (2004: 84), "stress was retracted to an immediately preceding syllable containing a vocalic element

followed by a laryngeal.” Note that this formulation excludes the sequence \*-VRH-: CS \*golvá ‘head’ and Lith. *galvà* preserve the final stress of \*golHúàH.

Baltic and Slavic also show agreement in accentual paradigms, the identity of which was established for nouns by Illič-Svityč (1963),<sup>7</sup> who derived the tripartite system established by Stang (1957) for Slavic<sup>8</sup> from its simpler Baltic counterpart (fixed-stress and mobile-stress paradigms). Illič-Svityč showed that Slavic accent paradigms (a) and (b) are in fact in complementary distribution and can be reduced to a single non-mobile paradigm, comparable to Baltic fixed stress. According to Illič-Svityč, stress is retained on an acute syllable (a.p. a), but advances from a non-acute (short or Balto-Slavic circumflex) syllable onto an immediately following syllable, resulting in a.p. b. Thus, while the root stress of the acute base Lith. *dūmai* ‘smoke’ corresponds to that of SCR. *dīm* (a.p. a), the root stress of the short base Lith. *blūsą* ‘flea, acc. sg.’, is matched by end stress in Russ. *блохы́* (a.p. b). Aside from the Slavic metatony associated with Meillet’s Law, Baltic and Slavic mobile bases retain their correspondence: Lith. *galvà* ~ *gálvą* ‘head, nom. sg. ~ acc. sg.’ (mobile a.p.) : Russ. *голова́* ~ *голове́* ‘head, nom. sg. ~ acc. sg.’ (a.p. c).

The Baltic bipartite paradigmatic accentual system (fixed and mobile) thus represents an earlier state from which the Slavic system can be derived. The assignment of a particular word to either the fixed or mobile class is, in Balto-Slavic terms, unmotivated. According to Vladimir Dybo (especially 1980: 147–150), the two accent classes result from a phonologization of earlier sequences of morphemes of either “dominant” or “recessive” valencies, making up a word form or phonetic word, with a “contour rule” according to which word stress falls on the first morpheme of dominant valency. If a word form (or phonetic word) consists only of recessive morphemes, the initial morpheme is automatically stressed. The valencies in this approach are seen as continuing a tonal opposition (high and low tone) inherited from Indo-European.

The derivation of the Slavic accentual model from an earlier Baltic model, noted above, suggests a lack of equivalence between the notions of Baltic and Slavic. Developing a point originally made by Endzelin (1911: 201–202), Stang (1966: 10ff.) notes that while Common Slavic presents a relatively monolithic system, Baltic is divided by a number of significant isoglosses, complicating the notion of Proto-Baltic. Certain of these isoglosses connect East Baltic (Lithuanian and Latvian) with a uniform Slavic, while others connect West Baltic (Old Prussian) with Slavic. The former include the Indo-European genitive singular \*-s(y)o of *o*-stem nouns, apparently preserved in Old Prussian (*deiwas* ‘God, gen. sg.’, and the pronominal *stesse* ‘that, gen. sg.’; cf. Slavic *česo* ‘what, gen. sg.’), but replaced in East Baltic and in Slavic by a form that has ablative function outside of Balto-Slavic: Lith. *rāgo* ‘horn, gen. sg.’, Latv. *raga* = OCS *roga* (East Baltic has extended this to the pronominal system as well: Lith. – *kō* ‘who, what, gen. sg.’, *tō* ‘that, gen. sg.’); a present passive participle in *-m* shared by East Baltic and Slavic (Lith. *vėdamas* ‘being led’, Latv. *vėdams* : OCS *vedomъ*), while Old Prussian continues an inherited *-manas* (in the hapax *poklausīmanas* ‘being heard’); East Baltic and Slavic pronouns have *-m* in the dative and locative singular, while Old Prussian continues inherited *-sm-* (Lith. *tām(ui)* ‘that, dat. sg.’, Slav. *tomu*; OPruss. *stesmu*); Lith. *devyni* ‘nine’, Latv. *deviņi*, Slav. *devětъ*, with an innovative *-d-*, while OPruss. *newīnts* ‘ninth’ preserves inherited *n-*. Finally, the word for ‘third’ reflects a stem \**tret-* in Lith. *trėčias*, Latv. *trėšs*, OCS *tretъjъ*, while OPruss. *tūrts* agrees with Skr. *tr̥tīyah*.

Isoglosses that link Old Prussian with Slavic include the possessive pronouns OPruss. *mais*, *twais*, *swais* ‘my, your, one’s own’, which agree with OCS *mojbъ*, *tvojbъ*, *svojbъ*, as opposed to Lith. (*manas*), *tavas*, *savas* and Latvian (*mans*), *tavs*, *savs*; the dative singular of the second person and reflexive pronouns preserves \*-b<sup>h</sup>- in Old Prussian (*tebbei*, *sebbei*) and Slavic (*tebě*, *sebě*) (Stang 1966: 248; Lithuanian and Latvian have rebuilt these forms: Lith. *táu*, *sáu*, Latv. *tev*, *sev*); the Old Prussian preterit of ‘to be’, *bēi*, *be* agrees with Slav. *běxъ*, *bě*, as opposed to the preterit stem \**bi-* found in Latv. dial. *biju* ‘(I) was’

and OLith. *bit(i)*; and the nasal-infixed Old Prussian participial forms *sindats* and *syndens* ‘sitting’ (Catechisms I and II) can be compared with the infixed CS \**sęd-* ‘sit’.

These isoglosses suggest a greater time depth for Common Baltic, “if indeed such a unified protolanguage ever was a reality” (Birnbaum 1970: 70–71). The resulting lack of equivalence between the notions of Baltic and Slavic (and the associated limitations of the family-tree model) is particularly emphasized by V. V. Ivanov and V. N. Toporov (1961). In reviewing the methodological preconditions for discussing the Balto-Slavic relationship, they argue that a relatively homogeneous Proto-Slavic can be derived from a considerably more archaic and heterogeneous Proto-Baltic linguistic model, in effect redefining the notion of Balto-Slavic by treating Slavic as a local development within a Baltic dialectal continuum.

Among more distant post-Indo-European linguistic affinities proposed for Balto-Slavic, correspondences with Germanic have been noted since Schleicher’s time (recall his “slawo-deutsch”). These include the marking of middle-intransitive voice through nasal infixation (a process that enjoyed particular productivity in Baltic) or nasal suffixation; the spread of the adjectival suffix \*-isko- in a common function: Goth. *-isks*, OCS *-ьskъ*, Lith. *-iškas* (Stang 1942: 275–278); and case forms in *-m-* that correspond to \*-b- elsewhere (for example, OLith. dat. pl. *-mus* (modern *-ms*), OCS *-mъ*, Goth. *-m*, as opposed to Skr. *-bhyah*, Lat. *-bus*: Meillet 1967: 148–152). Also characteristic are exclusive lexical agreements (Stang 1972: 79–82) across a number of semantic fields, including social relations: Lith. *draiūgas* ‘friend’, Latv. *drāugs*, OCS *drugъ*, Goth. *driugan* ‘to perform military service’; Lith. *liaudis* ‘people, nation’, OCS *ljudьje* ‘people’, OHG *liut*; Lith. *valdyti* ‘to govern, rule’ (3 pres. *valdo*), OCS *vladěti* ‘rule’, Goth. *waldan* || technical culture (simple tools, materials, handicrafts, objects of daily life): Lith. *aldijà* ‘dugout canoe’, OCS *aldii*, *ladii* ‘boat’, Norw. dial. *olda* ‘large trough, often from a carved-out tree trunk’; OPruss. *dalptan* ‘punch (tool)’, RussCS *dlato* ‘chisel’, Russ. *долбумъ* ‘to chisel’, OEng. *delfan* ‘to dig, delve’; Lith. *strėlė*, *strėlà* ‘arrow, shaft’, Latv. *strēla* ‘stripe, streak; jib, crane’, OCS *strěla* ‘arrow’, OHG *strāla*, OEng. *stræl* ‘arrow’ || products: Lith. *vāškas* ‘wax’, OCS *voskъ*, OEng. *weax*; Lith. *alūs* ‘beer’, ORuss., RussCS *olъ* ‘an intoxicating drink’, Sln. *ól* ‘beer’, ONor. *ól* ‘beer’ || agricultural terms: Lith. *rugiai* ‘rye’, OCS *ръжь*, ONor. *rugr* (*i*-stem) || certain adjectives: Lith. *sūrus* ‘salty’, OCS *syръ* ‘damp, moist’, ONor. *súrr* ‘sour’; Lith. *tingùs* ‘lazy, sluggish’, OCS *težьkъ* ‘heavy, difficult’, ONor. *þungr* ‘heavy’ || numbers: OPruss. *tūsimtons* ‘thousand’, Lith. *tūkstantis*, OCS *tysešti*, *tysošti*, Goth. *þusundi* (which do not fully agree in formation). Stang (1942: 278) suggests that these equivalences reflect a common preservation of inherited (perhaps northwest Indo-European) features, or represent convergent developments, and thus belong rather to the realm of Sprachbund phenomena.

## FURTHER READING

Henrik Birnbaum (1970) offers a balanced summary of various approaches to Balto-Slavic. Daniel Petit (2004) presents a historically oriented account of the Balto-Slavic question, from the Renaissance to the present day; he includes the findings of archaeology and provides an extensive list of references. Accentological issues are addressed by Rick Derksen (2004) in the same volume; an up-to-date overview of Balto-Slavic accentology can be found in the introduction to Derksen 2015. For a detailed presentation and evaluation of the data relating to Balto-Slavic, the careful and cautious studies of Christian S. Stang found in the references have not lost their relevance.

## NOTES

- 1 Baudouin de Courtenay (1903: 330–331) had earlier rejected a Balto-Slavic unity, calling it a “scholarly fiction” (учёная фикция), but without providing argumentation.



- 2 There are disagreements regarding the *-uR-* reflex of syllabic sonorants after labiovelars as a regular phonological change; see the discussion in Stang 1966 (77–82). Still, the fact of specific Baltic and Slavic agreements in the roots for ‘chase, drive’ and ‘drink, swallow’ is revealing. The *-i-* (rather than *-u-*) of Lith. *girtas* ‘drunk’, and the infinitives Lith. *giñti* ‘to drive (cattle)’, Latv. *dzīt*, is due to an extension of the productive ablaut pattern.
- 3 Derksen (2010) notes that this change was not shared by Old Prussian. But this argument depends on one’s interpretation of the orthographic evidence of the Old Prussian texts; see, for example, Schmalstieg 1976: 120.
- 4 Baltic characteristically preserves the original verbal root here, lacking in Slavic: Lith. *riñkti* ‘gather’ (3 pres. *reñka*), OPruss. *senrīnka*.
- 5 The role of accentology in the Balto-Slavic question was anticipated by Endzelin as early as 1911 (202): “дальнѣйшая разработка ударенія . . . , вѣроятно, еще болѣе освѣтитъ этотъ вопросъ” (further developments in stress . . . will undoubtedly shed more light on the issue).
- 6 There are a number of interpretations of Winter’s Law and its conditions of operation; for a review and critique of the literature, see Derksen 2002. Kortlandt (1988: 387) amends “lengthened vowels” in Winter’s original statement to “acute vowels,” since he rejects a connection between (non-laryngeal) length and acute tone. In Kortlandt’s view, Winter’s Law represents “the merger of the glottal element of the preglottalized voiced stop with the reflex of the Indo-European laryngeals, which had become a glottal stop in Balto-Slavic” (Derksen 2004: 83).
- 7 This progressive stress shift was demonstrated for Slavic verbs at approximately the same time by V. A. Dybo (1962), and is therefore referred to as either the Law of Illič-Svityč or Dybo’s Law.
- 8 Accentual paradigm (a.p.) a is characterized by columnar barytonic stress; a.p. b by columnar oxytonic stress; and a.p. c by final-initial mobile stress, which is compared with the oxytona of Greek and Vedic in thematic stems.

## REFERENCES

- Baudouin de Courtenay, Jan Niecisław 1903, “Lingvističeskie zametki i aforizmy, I-IV,” *Žurnal Ministerstva Narodnogo Prosvěščenija* 346: 279–334
- Birnbaum, Henrik 1970, “Four approaches to Balto-Slavic,” in: Rūķe-Draviņa, Velta (ed.), *Donum Balticum*, Almqvist & Wiksell, Stockholm: 69–76
- Brugmann, Karl 1904, *Kurze vergleichende Grammatik der indogermanischen Sprachen*, Karl J. Trübner, Strassburg
- Derksen, Rick 2002, “On the reception of Winter’s Law,” *Baltistica* 37/1: 5–13
- Derksen, Rick 2004, “Balto-Slavic accentuation: An update,” *Histoire Épistémologique Langage* 26/2: 81–92
- Derksen, Rick 2010, “The development of PIE \*eu in Baltic and Slavic,” in: Stadnik-Holzer, Elena & Holzer, Georg (eds.), *Sprache und Leben der frühmittelalterlichen Slaven. Festschrift für Radoslav Katičić zum 80. Geburtstag*, Peter Lang, Frankfurt am Main: 37–41
- Derksen, Rick 2015, *Etymological Dictionary of the Baltic Inherited Lexicon*, Brill, Leiden-Boston
- Dybo, V.A. 1962, “O rekonstrukcii udarenija v praslavjanskom glagole,” *Voprosy slavjanskogo jazykoznanija* 6: 3–27
- Dybo, V.A. 1980, “Balto-slavjanskaja akcentnaja sistema s tipologičeskoj točki zrenija i problema rekonstrukcii indoevropejskogo akcenta,” in: Sudnik, T.M. (ed.), *Balto-slavjanskije etnojazykovye kontakty*, Nauka, Moscow: 91–150

- Dybo, V.A. 2002, "Balto-Slavic accentology and Winter's Law," *Studia linguarum* 3/2: 295–515 [electronic version: <http://starling.rinet.ru/Texts/winter.pdf>]
- Endzelin, Jan 1911, *Slavjano-baltijskie étjudy*, M. Zil'berberg, Khar'kov
- Endzelin, Jan 1952, "Drevnejšie slavjano-baltijskie jazykovye svjazi," *Latvijas PSR Zinātņu akadēmijas Vēstis* III: 33–46
- Endzelīns, J. 1979, *Darbu izlase* III: 2, Zinātne, Riga: 429–445
- Hirt, Hermann 1892, "Vom schleifenden und gestossenen Ton in den indogermanischen Sprachen. I. Teil," *Indogermanische Forschungen* 1: 1–42
- Illič-Svityč, V.M. 1963, *Imennaja akcentuacija v baltijskom i slavjanskom: Sud'ba akcentuacionnyx paradigm*, Akademija nauk SSSR, Moscow [Illich-Svitych, V.M. 1979, *Nominal accentuation in Baltic and Slavic*, translated by R. L. Leed and R. F. Feldstein, MIT Press, Cambridge, Massachusetts]
- Ivanov, V.V. & Toporov, V.N. 1961, "K postanovke voprosa o drevnejšix otnošenijax baltijskix i slavjanskix jazykov," in: Tolstoj, N.I. (ed.), *Issledovanija po slavjanskomu jazykoznaniju* (Fourth International Congress of Slavists, Moscow, 1958), Izdatel'stvo Akademii nauk SSSR, Moscow: 273–305
- Kortlandt, Frederik 1985, "Long vowels in Balto-Slavic," *Baltistica* 21/2: 112–124
- Kortlandt, Frederik 1988, "Remarks on Winter's Law," in: Barentsen, A.A., Groen, B.M. & Springer, R. (eds.), *Dutch Contributions to the Tenth International Congress of Slavists, Sofia, Linguistics*, Rodopi, Amsterdam: 387–396
- Meillet, Antoine 1902, "O nekotoryx anomalijax udarenija v slavjanskix imenax," *Russkij filologičeskij vestnik* 48: 193–200
- Meillet, Antoine 1967, *The Indo-European Dialects*, translated by Samuel N. Rosenberg, University of Alabama Press, University, Alabama
- Otkupščikov, Ju. V. 1988, "Balto-slavjanskaja problema: Leksičeskij material i metody issledovanija," *Baltistica* 24/1: 11–26
- Petit, Daniel 2004, "Les langues baltiques et la question balto-slave," *Histoire Épistémologie Langage* 26/2: 7–41
- Schmalstieg, William R. 1976, *Studies in Old Prussian*, University Park, Pennsylvania: The Pennsylvania State University Press
- Smoczyński, Wojciech 2001, *Język litewski w perspektywie porównawczej*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków
- Stang, Christian S. 1942, *Das Slavische und Baltische Verbum*, Skrifter utgitt av Det Norske Videnskaps-Akademi i Oslo. II. Hist.-Filos. Klasse, no. 1, J. Dybwad, Oslo
- Stang, Christian S. 1957, *Slavonic Accentuation*, W. Nygaard, Oslo
- Stang, Christian S. 1966, *Vergleichende Grammatik der Baltischen Sprachen*, Universitetsforlaget, Oslo
- Stang, Christian S. 1970, "Bemerkungen über slavische und baltische Sprachen," in: Stang, Christian S. (ed.), *Opuscula linguistica*, Universitetsforlaget, Oslo: 53–64
- Stang, Christian S. 1972, *Lexikalische Sonderubereinstimmungen zwischen dem Slavischen, Baltischen und Germanischen*, Skrifter utgitt av det Norske Videnskaps-Akademi i Oslo, II. Hist.-filos. klasse, n. s., no. 11, Universitetsforlaget, Oslo
- Trautmann, Reinhold 1923, *Baltisch-Slavisches Wörterbuch*, Vandenhoeck & Ruprecht, Göttingen
- Vaillant, André 1950, *Grammaire comparée des langues slaves*, vol. 1, *Phonétique*, IAC, Lyon
- Winter, Werner 1976, "The distribution of short and long vowels in stems of the type Lith. *ėsti* : *vėsti* : *mėsti* and OCS *jasti* : *vesti* : *mesti* in Baltic and Slavic languages," in: Fisiak, Jacek (ed.), *Recent Developments in Historical Phonology*, Mouton, The Hague: 431–446

# BALTIC

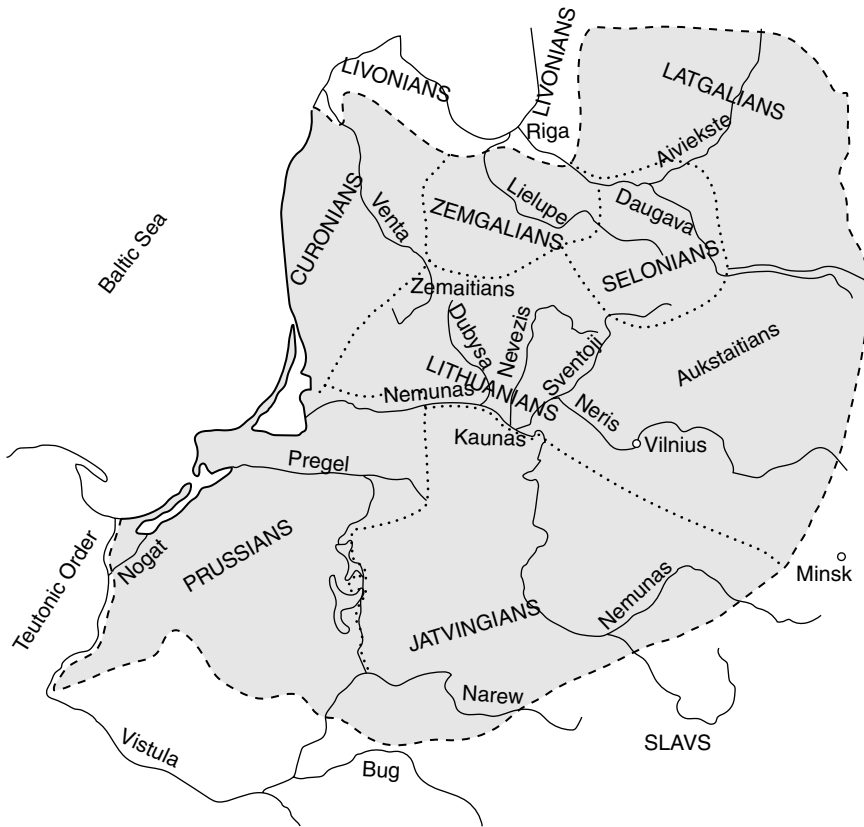
*Steven Young*

## INTRODUCTION

The Baltic language family is represented by Lithuanian and Latvian; a third major Baltic language, Old Prussian, became extinct by the beginning of the eighteenth century. Lithuanian and Latvian are traditionally grouped together as East Baltic, in contrast to Old Prussian (West Baltic); from a historical perspective, Lithuanian and Latvian may be viewed as representing core dialects within a Baltic (or Balto-Slavic) dialectal continuum, with Old Prussian as peripheral (Toporov 2006: 20–21). The original Baltic-speaking territory was once much larger, extending eastward into the upper Dniepr river basin and beyond.<sup>1</sup> These Baltic populations, among them the *Goljadz* (Balt. \*Galind-) of the eleventh- and twelfth-century Russian chronicles, were eventually assimilated by the eastern Slavs, a process continuing into modern times as former Lithuanian-speaking communities in what is now Belarus have become Slavicized. The paucity of surviving Baltic languages is somewhat compensated by the rich dialectal diversity within these languages; a number of dialectal features are undoubtedly due to the substratum influence of now-extinct Baltic languages – Curonian, Zemgalian, Jatvingian, and Selonian – about which little is known linguistically. An additional source of Baltic linguistic data are the many early (Bronze Age) Baltic borrowings into Baltic Finnic (see especially Kallio 2008) and Volga Finnic.

No common ethnonym for the ancestral Balts survives in the Baltic languages. The terms “Balt” and “Baltic” are secondary, after the name of the Baltic Sea (*mare Balticum* first appears in the eleventh century). The use of Baltic for the language family was first proposed in 1845 by Königsberg professor Georg H. F. Nesselmann in the introduction to his *Die Sprache der alten Preussen* (p. xxix). The lack of a common ethnographic designation is suggestive of the greater time depth and diffuse nature of the Baltic languages in comparison with the closely related (and more monolithic) Slavic language family; the reality of a Baltic protolanguage has even been questioned (Birnbäum 1970: 70–71). Lexical differences among the Baltic languages, and even between East Baltic Lithuanian and Latvian, can be quite striking. For example, Latvian and Lithuanian (together with Old Prussian) preserve different Indo-European words for ‘son’: Lith. *sūnūs*, OPruss. *soūns* (cf. OCS *synъ*, OInd. *sūnúh*), but Latv. *dēls*, related to Lat. *filius*; Lith. *kraujas* ‘blood’, OPruss. *crauygo*, *krawia* (cf. OCS *krъvъ*), but Latv. *asins*, an Indo-European *r/n* heteroclit cognate with OInd. *ásyṅ*, gen. sg. *asnáh*; Lith. *kárvė* ‘cow’ (cf. Russ. *koróva*), OPruss. *curwis* ‘ox’, but Latv. *gūovs* ‘cow’, continuing PIE \*g<sup>w</sup>eh<sub>2</sub>us.

Nevertheless, the notion of Common Baltic is justified by a number of characteristic innovations among the Baltic languages. These include a refashioned verb system, with the loss of a number distinction in the third person and the generalization of an *-a* (PIE \*-o) theme vowel in the present tense; a widespread nominal stem type in *-ē*; a number of common diminutive formations; an *-st-* formant in the present tense marking middle/intransitive meaning; and many characteristic lexical items, among them Lith. *lābas*



MAP 11.1 BALTIC TRIBES AT THE BEGINNING OF THE SECOND MILLENIUM AD

Source: *Lithuanian Encyclopedia*, Boston: Lithuanian Encyclopedia Press, Inc., vol. II, 1954: 148

‘good’, Latv., OPruss. *labs*; Lith. *ąžuolas* ‘oak’, Latv. *uõžuõls*, OPruss. *ansonis*; Lith. *briedis* ‘stag, elk’, Latv. *briõdis*, OPruss. *braydis*; Lith. *pelõ* ‘mouse’, Latv., OPruss. *pele*; Lith. *turõti* ‘to have’, Latv. *turõt*, OPruss. *turõt*; see Stang 1966: 7–9.

### Early texts and the emergence of standard languages

The oldest known text in a Baltic language is the Basel Epigram, two rhymed lines in Old Prussian from the mid-fourteenth century (see McCluskey, Schmalstieg, & Zeps 1975). The oldest more substantial text is the Elbing Vocabulary, a German–Old Prussian word list of 802 thematically arranged nouns and adjectives, making up the final seventeen pages of the Codex Neumannianus, which dates from about 1400 and is itself a copy of an original from the early fourteenth century (Schmalstieg 1976: 68). Another, rather inferior list of some one hundred German items with Old Prussian equivalents is found in Simon Grunau’s early sixteenth-century *Preussische Chronik*. A sense of the morphology

of Old Prussian is revealed in three sixteenth-century translations of Lutheran catechisms from German into Old Prussian; all suffer from a rather slavish translation from the German, and therefore shed little light on Old Prussian syntax. The first two, published in 1545 (the second is a “corrected” version of the first), each consist of sixteen pages, only about a third of which is actual Old Prussian text. The third catechism, published in 1561, includes German and Old Prussian versions of Martin Luther’s Enchiridion, or small catechism. The Old Prussian text makes up about 54 pages of the 134-page publication (Kabelka 1982: 54). The third catechism is our only source of Old Prussian word stress (and syllable tone).

The earliest monuments of Latvian and Lithuanian date only from the sixteenth century, when these languages already have a modern appearance. The earliest-known Lithuanian text is a handwritten set of prayers dating from the early 1500s. Book publication in Lithuanian began in East Prussia (which had a substantial Lithuanian population at the time), in connection with the spread of the Reformation. The first book published in Lithuanian is a 1547 translation of a Lutheran catechism by Martynas Mažvydas (Martinus Masvidius). The foreword begins with a personal appeal to the reader: “*Bralei seseris imkiet mani ir skaitikiet*” ‘Brothers, sisters, take me and read’. The language reflects Mažvydas’s native South Žemaitic dialect, with Aukštaitic elements. Subsequent East Prussian Lithuanian publications are written in an increasingly normalized variety of the local West Aukštaitic dialect, codified in Daniel Klein’s 1653 *Grammatica Litvanica*, the first grammar of Lithuanian. In the Catholic Grand Duchy of Lithuania, two writing traditions emerged, one based on the East Aukštaitic dialect of Vilnius, and the other representing the Central Aukštaitic dialect of the Kėdainiai region. The latter is reflected in Mykalojus Daukša’s 1595 translation of Jacobus Ledisma’s popular Catholic catechism, and his lengthy 1599 translation from the Polish of Jakub Wujek’s collection of sermons, the *Postilla Catholicka*. Daukša’s works are the first accented texts in Lithuanian, making them a valuable source for the study of Lithuanian historical prosody. The present-day standard language has its roots in the late nineteenth century and is based on the dialect of the southern West Aukštaitic region. Among the factors in the establishment of this variety as the national standard are the prior literary tradition of the virtually identical Aukštaitic dialect of neighboring East Prussia and the authority of the Lithuanian grammars of the Indo-Europeanist August Schleicher and Friedrich Kurschat, which described the same East Prussian Lithuanian speech.

The beginnings of a Latvian literary tradition also date from the Reformation. The first book published in Latvian is a 1585 translation of a Catholic catechism; this was soon followed by other religious texts, chiefly translations. The language of these early texts is to varying degrees influenced by the German speech of the authors, especially in syntax. Efforts to establish a national standard language were begun in the mid-nineteenth century; the Latvian traditional folksongs (*dainas*), which preserve a number of archaic features, served as an important source for norms. A milestone in the codification and description of modern Latvian was the appearance in 1922 of Jan Endzelin’s *Lettische Grammatik* and the four-volume *Latviešu valodas vārdnīca (Dictionary of Latvian)* (1923–32), begun by Karl Mühlēnbach (Kārlis Mīlēnbahs) and completed and edited by Jan Endzelin.

### East Baltic dialects

There are two major dialects of Lithuanian, the more conservative (at least in its southwestern subtype) Aukštaitic (*aukštaičių tarmė*), and Žemaitic (*žemaičių tarmė*; Samogitian), spoken in the northwestern quarter of Lithuania. Žemaitic presents a number of

structural and lexical similarities with Latvian, which may reflect the substratum influence of Curonian, an extinct Baltic language that was absorbed in the north by Latvian and in the south by Lithuanian.

The Latvian standard language is based on the central Latvian dialect (*vidus dialekts*), which is further divided (from west to east) into Curonian, Zemgalian, and Vidzeme varieties. The central dialect, together with Tamian in northwestern Curonia and along the northeastern coast of the Gulf of Riga, is known as Low Latvian. High Latvian (Selonian and Latgalian, the latter with an independent literary tradition; see Nau 2011), is found in the eastern third of the country. Latvian is traditionally viewed as representing a synthesis of the language of the early Latgalians (*Letthigalli*, *Lēt̃gola*, *Lotygola* of medieval chronicles), also known simply as Letts (*Letthi*, *Letti*), and neighboring closely related Baltic languages, now extinct, among them Zemgalian (along the Lielupe river), Curonian (in southwestern Latvia; a subtype of this dialect was spoken until the mid-twentieth century by the *kursenieki* of the Curonian Spit), and Selonian (along the middle Daugava river, extending into Lithuania).

## PHONOLOGY

### Baltic vowels

#### *The vowel system of Common Baltic*

The vowels of the Baltic languages can be derived from a Common Baltic system that, except for the merger of short \*a and \*o (shared by Germanic, Slavic, and Albanian) as \*a (or perhaps \*ɔ), reflects the vowels traditionally established for late Proto-Indo-European (for an alternative view, see the Kortlandt 1977 and 1985):

TABLE 11.1 THE VOWEL SYSTEM  
OF COMMON BALTIC

*i, *ī	*u, *ū
*e, *ē	*ō
*a, *ā	

The long vowels of Common Baltic ultimately reflect sequences of vowel plus laryngeal, the effects of Winter's Law (see below), and also (in the case of \*ē and \*ō) lengthened grade. In addition, we can establish a set of Common Baltic diphthongs consisting of \*e, \*a and a high vowel \*i, \*u (with a Balto-Slavic change of \*eu > \*īau, see p. 480), and a set of so-called mixed diphthongs, consisting of a vowel plus sonorant (\*r, \*l, \*m, \*n). The high-vowel mixed diphthongs continue Proto-Indo-European syllabic sonorants (see p. 479f.). The long vowels and diphthongs of Common Baltic were further distinguished (as they are in Common Slavic) by a prosodic opposition conventionally referred to as acute and circumflex; these structure points were subsequently realized, for the most part, as rising and falling contour tones in the daughter languages.

The Common Baltic vowel system is fairly well reflected in the Old Prussian of the Elbing Vocabulary (EV); a notable change is the raising and rounding of the low back vowels \*a (mainly after labials and velars) to \*ɔ (spelled <o>) and \*ā to \*ɔ̄ (spelled <o> or <oa>; Mažiulis 2004: 17). The Old Prussian catechisms, especially the third, reflect a number of innovations in the long vowels (to the extent that one can trust the orthography, which, after all, is the representation of a non-native speaker): the raising and rounding of

\*ā (and some instances of \*ō) to \*ū after velars and labials (*mūti* ‘mother’ < \*mātē : EV *mothe*; *mergu* ‘girl’ : EV *mergo*, Lith. *mergà*), the raising of \*ē to \*ī (*īst* ‘to eat’ : Lith. *ėsti*, Latv. *ēst* < \*ēd-; Mažiulis 2004: 16) and the (inconsistent) diphthongization of acute high vowels: *geīwan* ‘life, acc. sg.’, alongside *gīwan*, *gijwan*; *boūton*, *baūton* ‘to be’ alongside *būton*. Distinctive length was lost in unstressed syllables (Mažiulis 2004: 14).

### *The vowel system of East Baltic*

The somewhat skewed vowel system of Common Baltic, in which \*ō (with a close articulation: \*ō̄) lacks a short counterpart, emerges as a more balanced system in East Baltic. Here, the apparently quite open inherited pair \*e, \*ē was reinterpreted as the front counterpart of \*a, \*ā; the resulting gap in the system was filled by a new long front mid-vowel with a close articulation, \*ē̄ (parallel to \*ō̄), which developed under uncertain conditions from the diphthongs *-ej-* and *-aj-*, preserved in Old Prussian (for a detailed discussion, see Karaliūnas 1987: 152–179); compare Lith. *sniēgas* ‘snow’, Latv. *sniēgs* and OPruss. *snaygis*; Lith. *diēvas* ‘God’, Latv. *dievs* and OPruss. *deiwas*, as well as the Baltic Finnic borrowings Finn. *taivas* ‘sky, heaven’, Est. *taevas*; Finn. *heinä* ‘hay’ < CBalt. \*šeīna-, Lith. *šiėnas*. The new \*ē̄, together with its back counterpart \*ō̄, gave rise to the East Baltic “gliding” diphthongs, *ie*, *uo* (approximately [iə], [uə]).<sup>2</sup>

TABLE 11.2 THE VOWEL SYSTEM OF EAST BALTIC

*i, *ī	*u, *ū
*ē (> ie)	*ō̄ (> uo)
*e, *ē	*a, *ā

This realignment of the East Baltic system of long vowels was reflected in ablaut alternations, which were still productive at this period. Rather than the inherited alternation of an *ē*-grade opposing *ō*-grade, the productive pattern now becomes *ē* ~ *ā*, parallel to the *e* ~ *a* that continues Indo-European *e* ~ *o*; cf. Lith. *sodinti* ‘to plant’ (\*sād-), the causative to \*sēd- ‘to sit’. Lith. *sūodys* ‘soot’, Latv. *suōdrēji*, with old *ō*-grade \*sōd- to \*sēd- ‘to sit’, and Lith. *uōdas* ‘gnat, mosquito’, Latv. *uōds*, with old *ō*-grade \*ōd- to \*ēd- ‘to eat’, are isolated relic forms.

### *The vowel system of Lithuanian*

The above vowel system underlies later individual developments in Lithuanian and Latvian and their dialects. In both Lithuanian and Latvian, the high vowels and the short vowels have remained fairly faithful to the Common East Baltic system. Characteristic of the development of the Lithuanian vowel system was the repeated raising of long low vowels, with new low vowels emerging to take their place. The inherited long low \*ē and \*ā acquired a more close articulation and rose to fill the mid-vowel gaps created by the diphthongization of earlier \*ē̄ and \*ō̄: *ėjo* ‘went’ < \*ēīā (the new, close mid-vowel *ē* is spelled <ė>). The vacated slots were in turn filled by new long low vowels (the *e*, *ē* are quite open and can be represented as *ɛ*, *æ*), resulting from the denasalization of nasalized vowels; the latter arose from sequences of a short vowel plus tautosyllabic *n* before a continuant or in final position. Since this development post-dates the appearance of printed texts, the orthography reflects nasalization, denoted by a hook under the corresponding vowel character: *žąsis* [-a:-] ‘goose’, *tęsti* [-æ:-] ‘to continue’, *siųsti* [-ū:-] ‘to send’, *į*

[-ĩ:-] ‘into’. The functional load of the new  $\bar{a}$  and its  $\bar{e}$  counterpart was augmented by the results of a lengthening under stress (with concomitant circumflex tone) of most non-final short \*e and \*a; vowel length here is not indicated orthographically: *ledas* [lʲɛ̃ːdas] ‘ice’, *vakaras* [vãːkaras] ‘evening’ (compare Latv. *lēdus*, *vakars*, which preserve short e and a). There are exceptions to this general lengthening: infinitive roots (and derived forms) with a short low vowel remain short: *vēsti* ‘to lead’, *rāsti* ‘to find’, *vēsk!* ‘lead!’, *rāsk!* ‘find!’, *vēsiu* ‘(I) will lead’, *rāsi* ‘(you) will find’, as do the possessive pronouns *māno* ‘my’, *tavo* ‘your’, *savo* ‘his/her/its/one’s’, leading to the possibility of minimal pairs such as *māno* ‘thinks’: *māno* ‘my’.

After palatalized consonants and j, -a- and -ā- merge with -e-, -ē-, respectively: *giliàs* ‘deep, acc. pl. f.’, and *gilēs* ‘acorn, acc. pl.’, are homophonous: [gʲiˈlʲɛs]), as are *gilē* ‘acorn, acc. sg.’, and *giliq* ‘deep, acc. sg. f.’: [gʲiˈlʲæ:] (Girdenis 2003: 194).

The above changes result in the following vowel system for the Lithuanian standard language, represented in a broad phonetic transcription (corresponding graphemes are given in brackets):

TABLE 11.3 THE VOWEL SYSTEM OF LITHUANIAN

i <ɪ>, i: <y, ɪ̃>	u <u>, u: <ū, ũ>
e: <ē>	o: <ō>
ɛ <e, ia>, æ: <e, ia (under stress); ę, iq>	a <a>, a: <a (under stress); q>

This system has been rounded out by new short mid-vowel counterparts of  $\bar{e}$  and  $\bar{o}$ , but these (in particular ɛ) may be considered marginal phonemes, characteristic of borrowed words.

This inventory of Lithuanian vowels is supplemented by the “gliding” (*sutaptiniai*) diphthongs *ie* and *uo* (iə, uə), which arose from East Baltic \* $\bar{e}$  (< \*ei) and \* $\bar{o}$  (< \*ō) and function as long monophthongs, and by the pure diphthongs *au*, *ai*, *ei*, *ui*, and the mixed diphthongs, composed of a short vowel plus tautosyllabic sonorant: *i, u, e, a + r, l, m, n*.

#### The vowel system of Latvian

The Latvian vowel system generally reflects that of Common East Baltic, apart from a split in the quality of the low front vowels, originally conditioned by the nature of a following vowel: *e*,  $\bar{e}$  acquired a more close articulation before front vowels and palatal consonants; otherwise, they retained a more open articulation. Subsequent phonological and morphological processes rendered this split phonemic: *nesu* [nʲɛsu] ‘(I) carry’: *nesu* [nʲɛ̃su] ‘(I) carried’. The distinction between open and close *e*,  $\bar{e}$  is not reflected in Latvian orthography, although in linguistic texts the open *e*,  $\bar{e}$  is often denoted by ɛ, ē: the above “*nesu*” would thus be disambiguated as *nɛsu* [nʲɛsu] : *nesu* [nʲɛ̃su]. Long vowels are denoted in standard orthography with a macron. In linguistic works that include information on tone, the macron is replaced by one of the diacritics for tone; thus, the standard spelling *rīt* is a homograph for *rīt* ‘to swallow’ and *rīt* ‘tomorrow’.

TABLE 11.4 THE VOWEL SYSTEM OF LATVIAN

i, i:	u, u:
ɛ, ɛ:	
æ, æ:	a, a:



The functional load of the long high vowels  $\bar{i}$  and  $\bar{u}$  and the gliding diphthongs *ie*, *uo* (\* $\bar{e}$  < \* $e_i$ , \* $a_i$ ; \* $\bar{o}$ ) has been augmented by the results of a denasalization of the tautosyllabic sequences \*in, \*un, \*en, \*an, respectively: *mīt* ‘to tread’ : Lith. *minti*; *jūtu* ‘(I) feel’ : Lith. *juntū*; *pieci* ‘five’ : Lith. *penki*; *rūoka* ‘hand, arm’ : Lith. *rankà*. Occasional exceptions to this process are generally assumed to reflect substratum borrowings from the now-extinct Curonian, which did not undergo the change: *meñca* ‘cod’, *dziñtars* (alongside *džītars*) ‘amber’, *riñda* ‘row, line’.

The inventory of Latvian syllabics includes the pure diphthongs, among them *ei*, *ai*, *au*; the gliding diphthongs *ie* and *uo*, which function (as they do in Lithuanian) as long monophthongs (*uo* is spelled <o> in standard orthography but retained in linguistics texts, especially when syllable tones are marked), a practice adopted in this chapter: *rūoka* ‘hand, arm’, otherwise spelled *roka*); and the mixed diphthongs, composed of a short vowel plus tautosyllabic sonorant, although most instances involving tautosyllabic -n- will have become high vowels or gliding diphthongs (see above).

Characteristic of Latvian is the tendency to shorten final syllables in disyllabic and polysyllabic words.<sup>3</sup> Word-finally, all short vowels, with the exception of -u-, disappear: *vilks* ‘wolf’ : Lith. *vilkas*; *ass* ‘axle’ : Lith. *ašis*; *met* ‘throws’ : Lith. *mėta*; but (with -u-) *mėdus* ‘honey’ : Lith. *medūs*. Long vowels and diphthongs are shortened, with *ai*, *ei*, *ie* reduced to *i*, and *au*, *uo* to *u*: *rūokas* ‘hands, arms’ : Lith. *rañkos* (\*-ās); *lācis* ‘bear’ : Lith. *lokys* (\*-īs); *saki* ‘(you) say’ : Lith. *sakaĩ*; *saku* ‘(I) say’ : Lith. *sakaũ*; *mėdus* ‘honey, gen. sg.’ (thus homonymous with nom. sg. *mėdus*) : Lith. *medaũs*. Long vowels and diphthongs are preserved in monosyllabic forms: *tā* ‘that, gen. sg. m.’ : Lith. *tõ*, *tài* ‘that, dat. sg. f.’ : Lith. *taĩ*, *tiẽ* ‘that, nom. sg. m.’ : Lith. *tiẽ*.

## Prosodic structure

The prosodic evidence of the Baltic languages (together with that of Slavic) provides crucial information for reconstructing aspects of the phonological and accentual systems of Proto-Indo-European. In Common Baltic, as noted above, all long syllabics (long vowels or diphthongs, including sequences of vowel plus tautosyllabic sonorant) were characterized by a prosodic opposition of acute and circumflex, which have varied reflexes in the daughter languages.

### *Tonal oppositions in Lithuanian*

In standard Lithuanian, acute and circumflex are distinguished on stressed long vowels and diphthongs. Acute is realized as a falling tonal contour (in a mora analysis, the first mora is stressed), while circumflex is level or slightly rising (stress on the second mora of a long syllable). The tonal opposition is perceptually clearest on pure diphthongs: (acute) *šáu*! ‘shoot!’ : (circumflex) *šai*! ‘shout!’ The tones are indicated in reference works by diacritics: acute by ´ (placed over the first element of a diphthong) and circumflex by ˘ (placed over the second element of a diphthong). Short stressed vowels, which do not distinguish tone, are marked in reference works with a grave accent ( ` ); the grave is also used (in place of an acute diacritic) over a high vowel in an acute mixed diphthong: -*ir*-, -*ül*-.

The tonal contours of the standard language may be viewed as representing a point in a continuum of dialectal tonal variation. In most of Žemaitija, particularly the northwest, acute is a broken tone (*laužtinė priegaidė* [˘]), a rising-falling contour interrupted at its

peak by a glottal stop; in eastern Žemaitija and neighboring western Aukštaitic areas it weakens to the so-called Stosston (*stumtinė priegaidė*), which has a sudden rise followed by a falling contour. Further east, the acute loses its characteristic “Stoss” to become a simple falling tone. This is the variant of acute that has entered the linguistic literature (through the work of Friedrich Kurschat, who was the first to provide a description of the Lithuanian tones, based on his own East Prussian Lithuanian dialect) and is considered normative. Circumflex has a weak rising-falling contour in Žemaitija and western Aukštaitija, which, in moving east, “flattens out” on long monophthongs to simple length (protracted tone, *teštinė priegaidė*).

Unlike the Žemaitic dialects, where the peak of intensity for both acute and circumflex diphthongs occurs in the initial part of a syllable, Aukštaitic acute has its peak of intensity on the first element of a diphthong (which, if a low vowel, is lengthened: *kėlti* ‘to raise’ [kʲi̯e̯ːlʲtʲi]), while the circumflex reaches its peak on the second element of a diphthong, resulting in contour tones that are more stress- (or mora-) based than pitch-based; this is reflected in the Lithuanian terms for the tones: *tvirtapradė* (*priegaidė*) “strong-initial (tone)” for acute and *tvirtagālė* (*priegaidė*) “strong-final (tone)” for circumflex.

#### *Saussure’s Law and Leskien’s Law*

While the acute and circumflex tonal opposition in modern Lithuanian is distinguished only on stressed syllables, tones (or their prosodic antecedents) must once have characterized all long syllables, stressed and unstressed. This is shown by an earlier advancement of the ictus from a short or circumflex syllable to an adjacent acute syllable that had previously been unstressed: (nominal inflection) acc. sg. *rañka* ‘hand, arm’, but nom. sg. *rankà* (\*-ã); (nominal derivation) *kiaulė* ‘pig’ + *-ien-a* ‘meat of x’ > *kiauliena* ‘pork’; (verbal inflection) *lōšia* ‘plays’: *lošiù* (\*-úo) ‘(I) play’; (verbal derivation) *laiko* ‘holds, keeps’ + *-yti* (verb class marker, infinitive suffix) > *laikyti* ‘to hold, keep’. This progressive shift in ictus onto a following acute syllable was first recognized and described by Ferdinand de Saussure (Saussure 1896: 157) and is therefore known as Saussure’s Law.<sup>4</sup> In final syllables, the results of the operation of Saussure’s Law were obscured by a late post-dialectal process known as Leskien’s Law, according to which (for disyllabic and polysyllabic words) word-final acute syllables (including *ie* and *úo*) were shortened, forfeiting tone: \**ranká* > *rankà*. The gliding diphthongs *ie* and *úo* were replaced by *i* and *ù*; compare the long-form (definite) adjectives nom. pl. m. *gerieji* ‘good’, instr. sg. m. *gerúoju*, with the short forms *gerì*, *gerù*, which show the effects of Leskien’s Law. Diphthongs did not shorten but underwent metatony (change of tone) from acute to circumflex: \**matáu* > *mataũ* ‘(I) see’ \**matái* > *mataĩ* ‘(you) see’. The third-person future forms are exceptional in that non-high long monophthongs in both monosyllabic and polysyllabic words retain length, with circumflex metatony: *duōs* ‘give, 3 sg. fut.’: inf. *duoti*; *dēs* ‘put, 3 sg. fut.’: inf. *dėti*; *žinōs* ‘know, 3 sg. fut.’: inf. *žinoti*. Monosyllabic third-person future forms with acute high vowels were shortened: *būs* ‘be, 3 sg. fut.’: inf. *būti*; *lis* ‘rain, 3 sg. fut.’: inf. *lyti*.

As a result of Saussure’s Law, the fixed and mobile nominal accent paradigms inherited from Common Baltic each underwent a split into two paradigms, resulting in the four accent paradigms of modern Lithuanian: a.p. 1: fixed stress (non-stem-final ictus, or stem-final acute), a.p. 2: fixed stress with an overlay of Saussure’s Law (stem-final short or circumflex), a.p. 3: mobile stress (non-stem-final ictus, or stem-final acute), and a.p. 4: mobile stress plus an overlay of Saussure’s Law (stem-final short or circumflex).

TABLE 11.5 LITHUANIAN ACCENT PARADIGMS FOR A-STEM NOUNS

	a.p. 1	a.p. 2	a.p. 3	a.p. 4
<b>singular</b>				
nom.	výras ‘man’	piřstas ‘finger’	lángas ‘window’	vĩlkas ‘wolf’
gen.	výro	piřsto	lángo	vĩlko
dat.	výrui	piřstui	lángui	vĩlkui
acc.	výraq	piřstaq	lángaq	vĩlkaq
instr.	výru	piřstù	lángu	vĩlkù
loc.	výre	piřstè	lángè	vĩlkè
voc.	výre	piřšte	lángè	vĩlke
<b>plural</b>				
nom.	výrai	piřstai	langai	vilkaĩ
gen.	výru	piřstu	langũ	vilkũ
dat.	výrams	piřstams	langáms	vilkáms
acc.	výrus	piřstùs	lángus	vilkùs
instr.	výrais	piřstais	langaiš	vilkaĩš
loc.	výruose	piřstuose	languosè	vilkuosè

### Tonal oppositions in Latvian

In Latvian, tonal oppositions are found on (long) initial syllables, since initial stress has been generalized under the influence of neighboring Baltic Finnic languages (Stang 1966: 46). As in Lithuanian practice, the tones are indicated with diacritics in linguistic reference works; otherwise, they are not represented. Three phonemic tones are distinguished in the prosodically more conservative dialect of central Vidzeme: sustained (or weakly rising) tone (*stieptā intonācija*, denoted by ~, which in diphthongs is placed over the second element): *māte* ‘mother’ (Lith. *mótė*), *saule* ‘sun’ (Lith. *saulė*), *tīlts* ‘bridge’ (Lith. *tiltas*); broken tone<sup>5</sup> (*lauztā intonācija*, denoted by ^, which in diphthongs is placed over the second element): *rīts* ‘morning’ (Lith. *rýtas*), *galva* ‘head’ (Lith. *gálva*, acc. sg.), *būt* ‘to be’ (Lith. *būti*); and falling tone (*krītošā intonācija*, denoted by ` , which is placed over the first element of a diphthong): *rūoka* ‘hand, arm’ (Lith. *raĩnka*, acc. sg.), *draugs* ‘friend’ (Lith. *draũgas*), *pirkt* ‘to buy’ (Lith. *pirkti*). Tonal minimal pairs (or triplets) include *āust* ‘to dawn’ : *aũst* ‘to weave’; *rīt* ‘to swallow’ : *rīt* ‘tomorrow’; *tā* ‘that, nom. sg. f.’ : *tā* ‘that, gen. sg. m.’ : *tā* ‘thus’; *mīt* ‘to tread’ (Lith. *minti*), *mīt* ‘to change, exchange (arch., poetic)’ , *mīt* (inf. *mist*) ‘lives, dwells’ (Lith. *miĩta* ‘feeds on, nourishes oneself’).

Most Latvian dialect areas distinguish only two of these three tones. High Latvian has lost the distinction between falling and sustained tone in favor of falling; outside of High Latvian, broken and falling tone merge as either broken (in Kurzeme and Zemgale) or falling (Vidzeme), opposing sustained tone. The classic dictionary of Latvian, Mūlenbachs-Endzelĩns (Mūlenbachs 1923–32), uses a superscript <sup>2</sup> to indicate forms taken from dialect areas in which only two tones are distinguished.

### Endzelin’s Law

It follows from the above comparisons that Latvian sustained and broken tone both correspond to Lithuanian falling tone. In what has since become known as Endzelin’s Law, Jan Endzelin (1899: 267–268; 1922: 22) demonstrated that Latvian sustained tone corresponds to Lithuanian falling tone in a fixed-stress paradigm (a.p. 1): Latv. *brālis* ‘brother’ : Lith. *brólis*; Latv. *beržs* ‘birch’ : Lith. *béržas*, while Latvian broken tone corresponds to Lithuanian falling tone in a mobile paradigm (a.p. 3; that is, where there is an alternation between final and initial stress throughout the forms of a paradigm): Latv. *dzĩvs* ‘alive’ :

Lith. *gyvas*, *gyvą* (nom. sg. m./f.); Latv. *sirds* ‘heart’ : Lith. *širdis*, *širdį* (nom./acc. sg.). The remaining Latvian tone, falling, corresponds to Lithuanian rising tone (circumflex) in both fixed (a.p. 2) and mobile (a.p. 4) paradigms: Latv. *rūoka* ‘hand, arm’ : Lith. *rankà*, *rañka* (nom. sg., acc. sg.; a.p. 2); Latv. *zūoss* ‘goose’ : Lith. *žąsis*, *žąsį* (nom. sg., acc. sg.; a.p. 4); Latv. *drāugs* ‘friend’ : Lith. *draūgas* (a.p. 4); Latv. *džimt* ‘to be born’ : Lith. *gimti*.

A phonetic motivation for the two-fold reflex of Baltic acute in Latvian was first proposed by Richard Ekblom (1933: 69), who connected the development of Latvian broken tone with stress retraction onto an initial syllable. If this syllable was acute (which he assumed to be rising tone), it underwent an abrupt rise in pitch under new stress, leading to a “break” (“Umbruch”) in the syllable, represented by glottal closure and release. But an association between retracted stress and broken tone presents a number of difficulties and contradictions, which are avoided if broken tone, rather than sustained (or rising) tone, is taken to be primary; see Young 1994.

As a result of the distribution associated with Endzelin’s Law, Latvian occasionally preserves more information about the original fixed or mobile accent paradigm of an acute base than Lithuanian does. For example, the evidence of Latvian is needed in order to establish the original accentual class of disyllabic adjectives, since Lithuanian has generalized the mobile accent paradigm here: Latv. *balts* ‘white’, *ilgs* ‘long’, *pilns* ‘full’ indicate original fixed stress, while Lith. *báltas*, *-à*; *ilgas*, *-à*; *pilnas*, *-à* are secondarily mobile (a.p. 3). Latvian also continues an original fixed or mobile accent paradigm in infinitive forms, where Lithuanian provides no direct information: *aūgt* (Lith. *áugti*) ‘to grow’, *bēgt* (Lith. *bėgti*) ‘to run’, *duót* (Lith. *dúoti*) ‘to give’ point to original mobile stems, while *drāzt* (Lith. *drōžti*) ‘to carve’, *sēt* (Lith. *sėti*) ‘to sow’, *šāūt* (Lith. *šáuti*) ‘to shoot’ reflect original fixed stress.

### *Tonal oppositions in Old Prussian*

The Old Prussian Enchiridion also points to a tonal distinction between acute and circumflex. Assuming the traditional view that the macron that appears on vowel graphemes indirectly reflects stress (the macron marks long vowels, which were preserved only under stress), the representation of diphthongs with either the first or second element marked by a macron would suggest a falling or rising tonal contour (Fortunatov 1880: 153ff.; for a discussion of tones and tonal marking in Old Prussian, see Derksen 1998). Old Prussian examples reflecting Baltic circumflex include *ēit* ‘goes’ : Lith. *eīti* ‘to go’; *lāiku* ‘(they) hold to, keep’ : Lith. *laiko*; *āusins* ‘ear, acc. pl.’ : Latv. *āuss*, Lith. *āūsį* (acc. sg.). Among examples of Baltic acute are *kāulins* ‘bone, acc. pl.’ : Lith. *káulus* (acc. pl.); *pogaūt* ‘to receive, get’ : Lith. *pagáuti* ‘to catch’; *aīnan* ‘one, acc. sg. f.’ : Lith. *vieną* (acc. sg.). Moreover, acute (but not circumflex) *ī* and *ū* often diphthongize to *ei*, *ou* (*au*), allowing for placement of the macron on the second element of the diphthong (*eī*, *oū*): *boūt* ‘to be’ : Latv. *būt*, Lith. *būti*; *soūns* ‘son’ : Lith. *sūny* (acc. sg.); *geītwans* ‘alive, acc. sg.’ : Lith. *gyvas*. In the Elbing Vocabulary, a number of forms containing a diphthong that have circumflex counterparts in East Baltic show apparent lengthening of the first element, suggesting falling tone for circumflex: *doalgis* < \*dōlgis < \*dālgis ‘scythe’ : Lith. *dālgis*; *moasis* ‘bellows’ < \*mō(i)sis < \*mā(i)sis : Lith. *mašas* ‘bag’ (Mažiulis 2004: 14).

### *Interpretation of Baltic tonal correspondences*

We arrive at the following sets of tonal correspondences for the Baltic languages (the diphthong *au* here represents all long syllabics): Latv. *àu* (falling tone) = OPruss. *āu*

(falling tone) = Lith. *aũ* (rising tone) for Baltic circumflex; and Latv. *aũ/au̯* (rising or broken tone) = OPruss. *aī* (rising tone) = Lith. *áu* (falling tone) for Baltic acute. These equations, supported by the evidence of Slavic (as traditionally presented), led to the assumption of a falling pitch contour for circumflex and a rising pitch contour for acute in Proto-Baltic (and Balto-Slavic). Lithuanian was assumed to have undergone an inversion of tonal contours (see p. 481).

This statement of tonal correspondences has served as a traditional point of departure in discussions of Balto-Slavic tonogenesis. But acoustic studies by Aleksas Girdenis and Antanas Pakerys have shown that the Lithuanian tones (on diphthongs) differ more in the nature of the first component of the diphthong than in a prosodic contrast between the first and second components (Girdenis 2003: 288). Listening experiments using reverse recordings have shown that speakers of both Žemaitic and West Aukštaitic dialects of Lithuanian distinguish tones reproduced in the reversed direction just as well as original tones (Girdenis 2003: 273), which would not be expected if the primary exponent of tone was a rising or falling pitch contour. An acoustic characterization of the tones of prosodically conservative (North) Žemaitic Lithuanian is instructive here: for acute, acoustic energy is concentrated at a single point in the syllable nucleus and changes abruptly; circumflex is characterized by a lack of concentrated energy. A pitch contour in acute syllables can be explained as a side effect of the glottalization associated with broken tone (Vaillant 1936: 114–115, 1950: 244–245; Girdenis 2003: 272). This suggests that it is the broken (glottalized) tone of the Žemaitic Lithuanian dialects and Latvian that is the original reflex of Proto-Baltic (Balto-Slavic) acute.

## Baltic consonants

### *The Proto-Baltic consonant system*

The Proto-Baltic consonant system, which is essentially identical to that of Balto-Slavic (except perhaps for the Balto-Slavic reflex of the PIE palatovelars as *ś* and *ž*),<sup>6</sup> may be represented as follows:

TABLE 11.6 THE PROTO-BALTIC CONSONANT SYSTEM

	<i>labial</i>	<i>dental</i>	<i>alveopalatal</i>	<i>velar</i>
stops				
voiceless	p	t		k
voiced	b (PIE *b, b <sup>h</sup> )	d (PIE *d, d <sup>h</sup> )		g (PIE *g, g <sup>h</sup> , *g <sup>w</sup> , g <sup>wh</sup> )
fricatives				
voiceless		s	š	
voiced			ž (PIE *ǵ, ǵ <sup>h</sup> )	
nasals	m	n		
liquids		l, r		
glides	w		y	

Although the reflexes of Indo-European aspirated and plain voiced stops have merged as plain voiced segments, the original distribution is maintained in an adjacent syllabic: those preceding an original plain voiced stop are lengthened (if short monophthongs) and acquire acute tone. This process, known as Winter's Law, is illustrated by the following sets of correspondences:<sup>7</sup>

- PBalt. \*b < PIE \*b<sup>h</sup>, with circumflex base: Lith. *žẽmbti* ‘to cut, hew’, 3 sg. pres. *žẽmbia*: PIE \*ǵemb<sup>h</sup>- ‘to snatch, bite’, LIV<sup>2</sup> 162;
- PBalt. \*b < PIE \*b, with acute base: Lith. *obuolỹs* ‘apple’, acc. sg. *óbuolį*; Latv. *ábuõls*: PIE \*h<sub>2</sub>ebõl;
- PBalt. \*d < PIE \*d<sup>h</sup>, with short base: Lith. *vẽsti* ‘to lead’, 3 sg. pres. *vẽda*; Latv. *vẽd*: PIE \*wed<sup>h</sup>- ‘to lead’, LIV<sup>2</sup> 659;
- PBalt. \*d < PIE \*d, with acute base: *skiesti* ‘to dilute’, 3 sg. pres. *skiedžia*, Lith. dial. *skáidrús* ‘clear’; Latv. *skaĩdrs*: PIE \*sk<sup>h</sup>eyd- ‘to split, separate, tear’, LIV<sup>2</sup> 547–548;
- PBalt. \*g < PIE \*g<sup>h</sup> and \*g<sup>wh</sup>, with circumflex base: Lith. *algà* ‘wages, pay’, acc. sg. *álgą*; Latv. *àlga*; OPruss. gen. sg. *ālgas*: PIE \*h<sub>2</sub>elg<sup>wh</sup>- ‘to earn, yield (revenue)’, LIV<sup>2</sup> 263;
- PBalt. \*g < PIE \*g and \*g<sup>w</sup>, with acute base: Latv. *raũgs* ‘pupil (eye)’, *raũdzīt* ‘to see’: Gr. (Hesychius) ῥουγός : πρόσωπον;
- PBalt. \*ž < PIE \*ǵ<sup>h</sup>, with circumflex base: Lith. *liẽžti* ‘to lick’, 3 sg. pres. *liẽžia*; Latv. (with *o*-grade) *lāizīt*: PIE \*leyǵ<sup>h</sup>- ‘to lick’, LIV<sup>2</sup> 404;
- PBalt. \*ž < PIE \*ǵ, with acute base: Lith. *mélžti* ‘to milk’, 3 sg. pres. *mélžia*: PIE \*h<sub>2</sub>melǵ- ‘to milk’, LIV<sup>2</sup> 279.

### *Centum/satəm in Baltic*

While the Baltic reflexes of PIE \*k, \*ǵ, \*ǵ<sup>h</sup> are generally alveopalatal fricatives (a *satəm* development: Lith. *šimtas* ‘hundred’ < PIE \*k<sub>1</sub>m<sub>1</sub>tóm), there are occasional examples in which they continue as pure velars (a *centum* treatment). There are interesting correlations with Slavic, in which one language family or the other shows the exceptional *centum* development: Lith. *šeivà*, Latv. *saĩva* ‘bobbin’: CS \*koiuā in Russ. *чѣвка*; Lith. *šerti* ‘to feed’: CS \*kormъ ‘fodder’; Lith. *žvaizdẽ* (standard *žvaigždẽ*) ‘star’: CS \*gvězda; but with the opposite development in Lith. *klausyti* ‘to listen’, 3 sg. pres. *klaũso*: OCS *slyšati* ‘to hear’ (PIE \*k<sub>1</sub>lew<sup>s</sup>- ‘to hear, listen’, LIV<sup>2</sup> 336). Phonetic motivations have been sought for such exceptions, but these would still not account for lexical doublets such as Lith. *akmuõ* ‘stone’: *ašmuõ* (standard *ášmenys*, pl.) ‘blade’, Lith. *kiẽmas* (< \*keĩmas) ‘yard’: *šeimà* ‘family’, Lith. *kaũkti* ‘to howl’: *šaukti* ‘to shout’.

### *The RUKI rule in Baltic*

Another source for Proto-Baltic \*š is the historical process known as Pedersen’s Law or the *ruki* rule: the change of *-s-* to *-š-* after *k, r,* or the high vowels *ĩ, ũ*, operating in Baltic (unlike Slavic) before consonants as well as vowels. In Baltic, the change is regular after *-r-*: Lith. *viršus* ‘top’ (cf. ORuss. *врѣхъ*), Lith. *pirštas* ‘finger’ (cf. ORuss. *пѣръстѣ*, without the change). Reliable examples with *-k-* are difficult to find, although the frequent appearance of secondary *-kš-* makes it likely that the process took place here as well (Stang 1966: 96). Lithuanian generally preserves *-s-* after the high vowels: *sausas* ‘dry’: OCS *suxъ*; *ausis* ‘ear’: OCS *uxo*; *mūsos* ‘moss’: OCS *mъxъ*; *visas* ‘all’: WSlav. *vš-*, ORuss. (Novgorod) *врѣхъ*; *viesulas* ‘whirlwind’: RussCS *вѣхъ*, Lith. dial. *trisù* ‘three, loc.’: OCS *трѣхъ*, OInd. *trišù*. But there are a few examples that in fact reflect the rule: *aušrà* ‘dawn’ (but dial. *austrà* < \*ausrâ), *vetušas* ‘old (arch.)’: OCS *vetъxъ*, *maĩšas* ‘bag’: OCS *měxъ*, *jũšė* ‘fish soup’: Russ. *ухá*. This lack of consistency may result from Baltic being on the periphery for the diffusion of this areal feature (Stang 1966: 99).

*Later Baltic reflexes of the palatovelars*

Lithuanian alone among the Baltic languages preserves the reflexes *š* and *ž* of the Indo-European palatovelars and, for *š*, the output of the RUKI rule. In Latvian and Old Prussian (as in Slavic), these have become *s* (merging with inherited *s*) and *z*: Lith. *šuo* (PIE \**kwō-*) ‘dog’: Latv. *suns* (PIE \**kun-*); Lith. *šimtas* (PIE \**kmtó-*) ‘hundred’: Latv. *simts*; Lith. *žemė* (PIE \**d<sup>h</sup>ǵhem-*) ‘earth’: Latv. *zeme*, OPruss. *semē* (\**zemē*). But the alveopalatal fricatives must once have been more widespread, since early Baltic borrowings into Finnish (and other Baltic Finnic languages) show the reflex of \**š* and \**ž*, continuing in Finnish as *h*: Finn. *hammas* ‘tooth’: Lith. *žam̃bas* ‘sharp edge; sharp object’, Latv. *zūobs* ‘tooth’ (PIE \**ǵomb<sup>h</sup>-o-*); Finn. *herne* ‘pea’: Lith. *žirnis*, Latv. *ziņnis* (PIE \**ǵrHnó-* ‘grain’); Finn. *lohi* ‘salmon’: Lith. *lāšis*, *lašišà*, Latv. *lasis* (PIE \**loks-os-*).

*Sequences of consonant + \*j*

Proto-Baltic sequences of consonant plus \**j* before a back vowel result in various types of palatal or dental assimilations in the Baltic languages, which in turn give rise to morphophonemic alternations at stem boundaries. Where there is no assimilation, Latvian preserves the original *-j-* in the sequence, spelled <*j*>: *gulbja* ‘swan, gen. sg.’, while Lithuanian develops a palatalized consonant (orthographically, <*-Ci-*>), with secondary fronting of a low back vowel: *siūti* [‘s’iú:tʲi] ‘to sew’, *láukia* [‘lá:uk’i] ‘waits’, *labiaũ* [la‘biẽũ] ‘more’, *šiáurė* [‘š’á:ure:] ‘north’ (\**šjâur-* < \**šêur-* < \**keh<sub>1</sub>w(e)r-*, cf. CS \**sěver-*). The *j* in sequences with a labial stop is preserved in word-initial position in Lithuanian: *pjáuti* [‘pjá:utʲi] ‘to cut, mow’, *bjaurùs* [b’jœu‘ros] ‘ugly’. Latvian, like Slavic, introduces an epenthetic palatal *l* (spelled <*l̃*>): *plāūt*, *blāũrs* ‘angry, evil’.

Sequences of dental stop plus *j*, preserved in Old Prussian, perhaps as palatalized stops (cf. *median* ‘woods’: Lith. *mėdžias* ‘tree’, Latv. *mežs* ‘woods’), develop in Lithuanian into the alveopalatal affricates [tʃʲ, dʒʲ], orthographically <č(i)>, <dž(i)>: *čia* (\**tja*) ‘here’, *mėdžias* (\**medjās*) ‘woods’, *džiūti* (\**djũ-*) ‘to dry’. The corresponding reflexes in Latvian are the alveopalatal fricatives *š* and *ž*: *latvieša* (\**-tjā*) ‘Latvian, gen. sg.’; *mežs* (\**medjās*) ‘woods’: Lith. *mėdžias*; *žūt* (\**djũ-*) ‘to dry’: Lith. *džiūti*. The sequence \**s<sub>j</sub>* before a back vowel yielded *š* in Latvian: *šūt* (\**sjũt-*) ‘sew’: Lith. *siūti*, as it apparently did in Old Prussian: *schuwikis* ‘shoemaker’, cf. Lith. *siuvikis* ‘tailor’. Alveopalatal fricatives also result from \**s<sub>j</sub>* and \**z<sub>j</sub>* combinations in which the *s*, *z* continue the Indo-European palatovelars: *eža* (\**eziā* < \**ežjā*) ‘hedgehog, gen. sg.’ (nom. sg. *ezis*): Lith. *ėžis* (nom. sg. *ežys*).

Proto-Baltic sequences of velar stop plus *j* before a back vowel, as well as velar stop before a front vowel, resulted in fronted (palatalized) velars in Lithuanian, but assibilated to dental affricates (orthographically, <*c*>, <*dz*>) in Latvian. Sequences of velar stop plus *j*: Latv. *caũne*: Lith. *kiáunė* ‘marten’, *acu* ‘eye, gen. pl.’: Lith. *akiũ*; *rudzi* ‘rye, nom. pl.’: Lith. *rugiaĩ*; velar stop before a front vowel: Latv. *acs* ‘eye’: Lith. *akis*, Latv. *lācis* ‘bear’: Lith. *lokys*, Latv. *dzimt* ‘to be born’: Lith. *gimti*. A secondary palatalization of the new dental affricates can be found in masculine *-ja*-stems, where a stem-final *-c-* or *-dz-* merged with a restored *-j-* to produce a new *č*, *dž*: *lāča* (< \**lācjā*) ‘bear, gen. sg.’ (nom. sg. *lācs*): Lith. *lōkio* (nom. sg. *lokys*). Note also Latv. *čāula* ‘hull, husk; shell’ compared to Lith. *kėvalas* ‘shell’; its historical development serves as a recapitulation of some of the sound changes presented here: \**kėvala* > \**cėvala* (assibilated of a velar stop before a front vowel) > \**cėula* (vowel syncopation and concomitant formation of a falling-tone diphthong) > \**cjāula* (\**eu* > \**iau* as a persistent rule in East Baltic, dating from the Balto-Slavic period) > *čāula* (palatalization of *cj* to *č*).

*Dissimilation in clusters of dental stops*

Sequences of dental stops continue in Baltic (as in Slavic) as fricative plus stop clusters: Lith. *mès-ti* (3 sg. pres. *mēt-a*), past pass. ptcp. *mès-tas* (\**met-t-*) ‘to throw’ = Latv. *mes-t* (3 sg. pres. *mēt*), past pass. ptcp. *mes-ts*; Lith. *vès-ti* (3 sg. pres. *vēd-a*), past pass. ptcp. *vès-tas* (\**wed-t-*) ‘to lead’ = Latv. *ves-t* (3 sg. pres. *vēd*), past pass. ptcp. *ves-ts* = OPruss. *wes-t*.

*Dissimilation in clusters of dental stop + l*

A dissimilation of the dental in the clusters *-tl-*, *-dl-* to a velar (*-kl-*, *-gl-*) is found in East Baltic and, in part, in Old Prussian: Lith. *žénklas* ‘sign’ (with *-kla-* < \**-tla-*), but OPruss. *ebsentliuns* (\**-zen-tl-*) ‘designated’; Lith. *ėglė* ‘spruce, fir’, Latv. *egle*, but OPruss. (Elbing Vocabulary) *addle*, where the dental is preserved. The Old Prussian of the Elbing Vocabulary is inconsistent here; despite *addle* ‘spruce, fir’, the dissimilatory change appears in *clockis* ‘bear’ (Lith. *lokj̄s*, Latv. *lācis*, CBalt. \**tlōk-*). The unchanged sequence is preserved in the Old Prussian place name *Tlokun=pelek* ‘bear swamp’, the first element of which is the genitive plural of ‘bear’.

**MORPHOLOGY****The noun***Nominal inflection: case, number, gender*

The Baltic languages inherited from late Indo-European a system of cases for marking the syntactic and semantic role of a noun phrase in a sentence: nominative, genitive (continuing, as in Slavic, both genitive and ablative function), dative, accusative, and instrumental, together with a vocative form of address. A locative case was also inherited (PIE \**-oy*, cf. adverbialized *o*-stems Lith. *namiē* ‘at home’, OPruss. *bītai* ‘in the evening’; *ā*-stem *rañkoje* < \**rank-oi-en* ‘in the hand’), but was rebuilt in East Baltic as part of a system of secondary local cases formed with postpositions added to existing case forms. In addition to the inessive (old locative + \**en*; Stang 1966: 182), this system included an illative (accusative + \**n(a)*), adessive (old locative + \**p(i)*), and allative (genitive + \**p(i)*). Besides the standard inessive (simple locative), certain of these, particularly the illative, are still found dialectally in Lithuanian, or continue as adverbs: *laukañ* ‘outside (directional)’. It is difficult to establish in a reliable way the case system of Old Prussian, due to possible morphological interference from German in the texts; it undoubtedly agreed with East Baltic in marking at least a nominative, accusative, genitive, dative, and vocative, and possibly a locative (and, among personal pronouns, an instrumental).

In the modern Baltic languages, the various cases are distinguished in two numbers, the singular and plural; a dual is attested for certain case forms in Lithuanian dialects and in older Lithuanian and Latvian texts. The noun has inherent masculine or feminine gender; the Balto-Slavic neuter merged in East Baltic with the masculine: Lith. *šiėnas* ‘hay’ (m.), Latv. *siens* (m.) : OCS *sěno* (n.). The neuter, with a nominative-accusative *a*-stem ending *-an*, is attested in the Old Prussian of the Elbing Vocabulary, for example, *assaran* (\**azaran*) ‘lake’ (East Baltic masculine Lith. *ėžeras*, Latv. *ezers*) and the phrase *ructan dadan* (spelled as a single word) ‘sour milk’, cf. Latv. *rūgts* ‘bitter’; the *u*-stem *alu* ‘mead’ (: Lith. *alus*, Latv. *alus* ‘beer’) shows a zero ending. By the time of the catechisms, the Old Prussian neuter was in the process of disappearing (Petit 2010: 153).



*Nominal stem types*

The endings marking the grammatical categories of case, number, and gender follow several inherited declensional patterns, based on stem type; these are most clearly preserved in Lithuanian and give the Lithuanian noun its characteristic archaic appearance. The following stem types are represented in Lithuanian; Latvian has a somewhat reduced and simplified system:

- **(j)a-stems**, which continue Indo-European (thematic) *o*-stems and include masculines and neuters (which merged in East Baltic with masculines): (masculine) Lith. *vilkas* ‘wolf’, Latv. *vilkš* < PIE \*w<sub>l</sub>kʷos; OPruss. *deiws* ~ *deiwas* (1x) ~ *deiwis* (EV) ‘God’; (original neuter) Lith. *kāulas* ‘bone’, cf. OPruss. (EV) *caulan* < PIE \*keh<sub>2</sub>ulo-, Lith. *būtas* ‘lodgings’: OPruss. *buttan*; *ja*-stem: Lith. *svėčias* ‘guest’, Latv. *svešs* ‘guest; strange, unknown’ < EBalt. \*svetjas; the Latvian consonantal stem *zvērs* ‘beast’ has joined the *a*-stems, apparently on the basis of the genitive plural *zvēru* (Gr. θηρῶν);
- **ī-stems**, which develop from the sequence *-ija-*: Lith. *brōlis* ‘brother’, Latv. *brālis* (EBalt. \*brālīs, PBalt. \*brālijas); Lith. *dagys* ‘thistle’, Latv. *dadzis* ‘thistle, burdock’ (\*dagijas; the original *-ija-* sequence is seen in the Estonian borrowing *takijas* ‘thistle, burdock’); the *-ys* of Lithuanian represents a contraction of \**-ijās*. This type also includes some masculine consonantal stems: Lith. *mėnuo*, gen. sg. *mėnesio*, nom. pl. *mėnesiai* ‘moon, month’, Latv. *mēnes(i)s*, gen. sg. *mēneša*, nom. pl. *mēneši* ‘moon, month’ < PIE \*meh<sub>1</sub>n(e)s); Latv. *suns*, gen. sg. *suņa*, nom. pl. *suņi* ‘dog’; Latv. *ūdens*, gen. sg. *ūdens*, acc. sg. *ūdeni*, nom. pl. *ūdeņi* ‘water’; Latv. *akmens*, gen. sg. *akmens*, acc. sg. *akmeni*, nom. pl. *akmeņi* ‘stone’;
- **(j)ā-stems**, continuing PIE \*eh<sub>2</sub>-stems: Lith. *rasà* ‘dew’, Latv. *rasa* < PIE \*roseh<sub>2</sub>; *iā*-stem: Lith. *žinià* ‘(piece of) news’, Latv. *ziņa*. This class also includes the Lithuanian relic forms *marti*, acc. sg. *marčią* ‘daughter-in-law’ and *pati*, acc. sg. *pāčią* ‘wife’, and the feminine forms of the present active participles: *dirbanti* ‘working’, acc. sg. *dirbančią*, which continue the type represented by Skr. *devī* ‘goddess’;
- **ē-stems**, most of which represent a development from \**-ijā* (Stang 1966: 201–204): *saulė* ‘sun’, Latv. *saule* < CBalt. \*sāulijā, in Indo-European terms \*seh<sub>2</sub>ulijeh<sub>2</sub>; *žėmė* ‘earth’, Latv. *zeme*, OPruss. *semmē* < CBalt. \*žemijā, comparable to Slav. \*zemjā. The development of \**-ijā* to *-ē* must have been early, since the latter is already reflected in borrowings into Finnish: *kantele* ‘a kind of stringed instrument’: Lith. *kañklės* (\*kañtlės);
- **i-stems**, continuing Indo-European *i*-stems, most of which are feminine (Lith. *avis* ‘sheep’, Latv. (arch.) *avs* < PIE \*h<sub>3</sub>ewis); Lith. *ugnis* (OLith., ELith. *ùgnis*) ‘fire’, Latv. *uguns*, now feminine, is attested as masculine in Old Lithuanian (cf. OCS *ognь*, m.; secondarily *jo*-stem *ognь*). The inherited *i*-stem declension has absorbed various consonantal stem types; the entry point for the consonantal stems was undoubtedly the accusative, in which the Balto-Slavic reflex of the syllabic nasal case marker in consonant stems (e.g., PIE \*h<sub>3</sub>dont-ṛ̥ ‘tooth, acc. sg.’ > CBalt. \*dant-in, Lith. *dañtį*) merged with the ending of an *i*-stem, for example, \**avi-n* ‘sheep’. Old inherited feminine consonantal stems include Lith. *naktis*, acc. sg. *naktį* ‘night’, Latv. *nakts* < PIE \*nokʷts; Lith. *žąsis*, acc. sg. *žąsį* ‘goose’, Latv. *zoss* < PIE \*ǵʰ<sub>2</sub>ens; Lith. *širdis*, acc. sg. *širdį* ‘heart’ < PIE \*k<sub>1</sub>rd-; Lith. *žuvis*, acc. sg. *žuvį* ‘fish’ < CBalt. \*žū (PIE \*dǵʰuH), acc. sg. \*žuṽin; Lith. *ausis*, acc. sg. *aušį*, Latv. *auss* < PIE \*h<sub>2</sub>ews; and Lith. *duktė*, acc. sg. *dukterį* ‘daughter’ < PIE \*dʰugh<sub>2</sub>tēr. Examples of masculine

consonantal stems are Lith. *dantis*, gen. sg. *dantiės* ‘tooth’ < PIE \*h<sub>3</sub>dont(s); Lith. *žvėris*, gen. sg. *žvėriės* ‘beast’, cf. Gr. θήρ; Lith. *šuō*, acc. sg. *šūnį*, Latv. *suns* < PBalt. \*š(u)ō, \*šun-, PIE \*ku > w ōn; Lith. *vanduō*, dial. *vānduo*, acc. sg. *vāndenį*, Žemaitic Lith. *unduō*, *ūnduo*, Latv. *ūdens*, OPruss. *wundan*, *unds* < CBalt. \*uāndō, \*ūnden-; and stems with the extended formant *-men-*, which in Baltic generalize a nom. sg. *-mō(n)* and an oblique stem *-men-*: Lith. *akmuō*, acc. sg. *ākmenį* ‘stone’, Latv. *akmens* < PIE \*h<sub>2</sub>ekmōn; Lith. *piemuō*, acc. sg. *piemenį* ‘shepherd’, compare Gr. ποιμήν. In Latvian, where the *i*-stems are exclusively feminine, the masculine consonantal stems have mostly merged with the *ī*-stems, sometimes preserving the original consonantal genitive singular: *akmens* (\*akmen-es) ‘stone’, *ūdens* (\*uānden-es) ‘water’. Vestiges of the Indo-European consonantal declension are preserved in Lithuanian (standard language, dialects, old texts) in gen. sg. *dukterīs* (~ *dukterės*), *akmeīs* (~ *akmenės*), *šūnīs* (~ *šūnės*); nom. pl. (dial.) *ākmenes* ‘stone’, *dūkteres* ‘daughter’, *žvėres* ‘beast’ (Gr. θήρες), *žūves* ‘fish’ (Gr. ἰχθύες), *žq̄ses* ‘goose’ (Gr. χῆνες); and gen. pl. *dantū* to *dantis* ‘tooth’ (Gr. ὀδόντων), *naktū* to *naktis* ‘night’ (Gr. νυκτών), *dukterū* to *duktė* ‘daughter’, *žuvū* to *žuvīs* ‘fish’ (Gr. ἰχθύων), and *žqsū* to *žqsīs* ‘goose’ (Gr. χηνῶν). Latvian also preserves a number of such genitive plural forms dialectally (Endzelin 1922: 318–319): *āusu* to *āuss* (Lith. *ausis*) ‘ear’ (PIE \*h<sub>2</sub>ews), *zūosu* to *zūoss* (Lith. *žqsīs*) ‘goose’, *naktu* (standard *nakšu*, an *i*-stem) to *nakts* (Lith. *naktis*) ‘night’, *zuvu* to *zivs* (standard *zivs*, gen. pl. *zivju*) ‘fish’;

- ***u*-stems**, continuing Indo-European *u*-stems (Lith. *sūnūs* ‘son’ < PIE \*suHnus; Lith. *ledūs* ‘ice’, Latv. *lēdus*, compare OCS *ledъ*); most are masculine, but they include a few old neuters: Lith. dial. *pēkus* ‘livestock’, OPruss. *pecku*; Lith. *alus* ‘beer’, Latv. *alus*, OPruss. *alu*. Latvian preserves, as pluralia tantum, a few old \*ū- (\*uH-) stem feminines, such as *dziirņus* (also *dziirņavas*) ‘millstone’ (PIE \*g<sup>w</sup>ṛh<sub>2</sub>nuHs; compare Slav. (OCS) *žръны*).

The more productive stem types have become strongly associated with a particular gender: (*ī*)*a*- and *ī*- (*ija*-) stems, together with *u*-stems, are masculine; (*ī*)*ā*- and *ē*- (*ijā*-) stems, together with *i*-stems, are feminine. There is a strong tendency in the Lithuanian dialects for the few masculine *i*-stems, such as *žvėris*, to be treated as feminine. As we have seen, in Latvian the old masculine *i*-stem *zvērs* has joined the *ī*-stem masculines.

### Sample *o*-stem case endings

What follows is a brief sketch of individual case endings for *o*-stem (Baltic *a*-stem) nouns. PIE nom. sg. \*-s, associated in Baltic with masculine gender, was extended in East Baltic to neuters, as these merged with the masculines: Lith. *ėžeras* ‘lake’: OPruss. *assarān* (\*azarān), with *-n* < PIE \*-m. In East Baltic, as in Slavic, nouns in the genitive assumed the original ablative ending \*-o-Hed, resulting in a circumflex \*-ā: Lith. *vilko* ‘wolf’, Latv. *vilka*. The Old Prussian genitive apparently preserves an original *-s-* formant: *deiwas* ‘of God’. Lithuanian dative *-ui* (circumflex) continues PIE \*-o-Hei, through \*-uoi < CBalt. \*-ōi. The accusative singular (and plural) are formed with the addition of *-n* (PIE \*-m) to the theme vowel: Lith. *vilką*. The instrumental adds a laryngeal to the theme vowel, resulting in an acute CBalt. \*-ô: Lith. *vilkiu*. The Old Prussian adverbial *bītai* ‘in the evening’ preserves an old locative ending *-ai* (PIE \*-o-y), also found as a fossilized form in the Lithuanian adverb *namiē* ‘at home’. The regular ending

seen in Lith. *vilķē* ‘in the wolf’ (Latvian has generalized the *ā*-stem ending) seems to reflect an inessive *-en* added to the old locative form, with an unmotivated acute. The Baltic languages preserve, in the singular, an inherited vocative with the bare theme vowel *-e*: OPruss. *O Deīwe* ‘O God!’, *Tāwa* (~ *Tawe*) *Noūson* ‘Our Father’; Latv. *tēvs* < \**tēvs* (modern *tēvs*) ‘father!’; Lith. *Diēve* ‘God!’, *Tēve mūsu* ‘Our Father’, *vilķe* ‘O wolf!’; for personal names, this formation has been replaced in Lithuanian by the particle *-ai*: *Jōnai!* (nom. *Jōnas* ‘John’).

The nominative-accusative dual ending in (*du*) *vilķū* continues an Indo-European laryngeal added to the thematic stem, while the dative dual *vilķām* and instrumental *vilķaṃ* go back to \*-a-ma-. Nominative plural *-ai* (Lith. *vilķai*), also found in Old Prussian (*wijrai* ‘men’), is undoubtedly borrowed from the pronominal stems. A neuter nominative plural in *-ā* is attested in the Old Prussian Elbing Vocabulary: *warto* ‘door’ (cf. OCS *vrata* ‘gate, nom./acc. pl.’). The genitive plural ending of Lith. *vilķū*, and the Old Prussian place name *Tlokunpelk* ‘bear swamp’, represents a Proto-Baltic circumflex \*-ōn (> EBalt. \*-uon > *-ū*), continuing PIE \*-o-om; the Catechism texts often show *-an*, which would reflect a short \*-on. The earlier Lithuanian dative ending *-mus*, shortened to *-ms* (*vilķams*), does not fully agree with OPruss. *-mans* (*waikammans* ‘servant, dat. pl.’), but see Mažiulis 2004: 43. The East Baltic accusative plural (Lith. *vilķūs*) apparently represents \*-ōns, with inorganic length and acute tone; Old Prussian has an expected *-ans*: *deiwan* ‘gods’. The instrumental plural of Lith. *vilķais* represents a PIE \*-o-Heys. The locative (inessive) plural seen in *vilķuosē* is secondary; according to Stang 1966: 186, the original ending is preserved in adverbial numerals such as *keturīese* ‘four (together)’, dial. *keturīesu* (Zinkevičius 1966: 324), continuing BSl. \*-oysu; cf. Slav. (OCS) *vlbčěxъ* ‘wolf, loc. pl.’ Sample declensions of Baltic *a*-stem nouns in the three Baltic languages are given in table 11.4.

TABLE 11.7 DECLENSION OF A-STEM NOUNS

	Lithuanian	Latvian	Old Prussian
<b>singular</b>			
nom.	<i>vilķas</i> ‘wolf’ (a.p. 4)	<i>vilķs</i> ‘wolf’	* <i>grīks</i> ‘sin’, cf. <i>waix</i> (* <i>vaiks</i> ) ‘servant’
gen.	<i>vilķo</i>	<i>vilķa</i>	<i>grīkas</i>
dat.	<i>vilķui</i>	<i>vilķam</i>	<i>grīku</i>
acc.	<i>vilķą</i>	<i>vilķu</i>	<i>grīkan</i>
instr.	<i>vilķū</i>	= acc.	?
loc.	<i>vilķē</i>	<i>vilķā</i>	* <i>grīkai</i> , cf. (adverbial) <i>bītai</i> ‘in the evening’
voc.	<i>vilķe</i>	<i>vilķ(s)</i>	* <i>grīke</i> , cf. <i>deīwe</i> ‘God’
<b>dual</b>			
nom.-acc.	<i>vilķū</i>		
dat.	<i>vilķām</i>		
instr.	<i>vilķaṃ</i>		
<b>plural</b>			
nom.	<i>vilķai</i>	<i>vilķi</i>	<i>grīkai</i>
gen.	<i>vilķų</i>	<i>vilķu</i>	<i>grīkan</i>
dat.	<i>vilķams</i>	<i>vilķiēm</i>	* <i>grīkamans</i> , cf. <i>waikammans</i> ‘servant’
acc.	<i>vilķūs</i>	<i>vilķus</i>	<i>grīkans</i>
instr.	<i>vilķais</i>	= dat.	?
loc.	<i>vilķuosē</i>	<i>vilķuōs</i> (sp.: <i>vilķos</i> )	?

### Nominal derivation

Among the older derivational formations are deverbal agent nouns with an \*-āja- suffix added to the infinitive stem: Lith. *giedotojas* ‘choir-boy’ (inf. *giedoti* ‘to sing, chant’, Latv. *dziēdātājs* ‘singer’ (inf. *dziēdāt* ‘to sing’), with traces in Slavic: OCS *ratajъ* ‘plowman’ : OPruss. *artoyš*, Lith. *artojas* (inf. *arti* ‘to plow’). Verbal noun formations differ across Baltic: Lithuanian uses *-imas*, added to the preterit stem: *piešimas* ‘drawing’ (*piešti* ‘to draw’); the latter has a direct cognate in CS \*рiсьмо ‘writing’ (Pol. *pismo*, Cz. *písmo*). In Latvian, verbal nouns are generally formed with the suffix *-šana* (Mathiassen 1997: 158): *rakstīšana* : *rakstīt* ‘to write’, while in Lithuanian this suffix refers to the manner in which something is done: *rašysena* ‘handwriting’ (i.e., how one writes). The corresponding suffix in Old Prussian is *-snā*: *rickaūsnan* ‘government’. A widespread suffix, varying somewhat in form, denotes persons according to their occupation, place of residence, or some characteristic: Lith. *darbiniškas* ‘worker’ (*dárbas* ‘work’), Latv. *darbiniēks* ‘employee, office-worker’ (*darbs* ‘work’), OPruss. *balgninix* ‘saddler’ (*balgnan* ‘saddle’); the Slavic cognate *-ьникъ* again demonstrates the closeness of the Baltic and Slavic language families.

### The adjective

The grammatical categories of case, number, and gender (masculine or feminine), inherent in the noun, are reflected in an associated adjective. Thematic neuter adjectives are attested in the Old Prussian Elbing Vocabulary (*kirsnan* ‘black’, cf. Lith. *kišnas* ‘jet-black (horse)’, CS \*čъrnъ ‘black’, and Skr. *कृष्णः* ‘black’) and in the Enchiridion: *nawnan* ‘new’; the third catechism also has the neuter *u*-stem (used as an adverb) *polīgu* ‘like’ and the *i*-stem *arwi* ‘true’. A vestigial adjectival neuter can be found in Lithuanian in predicate adjectives and certain pronominal forms (Petit 2010: 171–174).

### Declension types for simple adjectives

The declensional paradigms of the adjective reflect those of the noun, with the substitution in Lithuanian of pronominal endings in the dative singular and plural masculine (*-ám*, *-iems*), the locative singular masculine (*-amè*), and, for (*i*)*a*-stems, the nominative plural masculine (*-i*). The Latvian adjective follows the corresponding noun declension, which already incorporates pronominal endings in the dative singular and plural for masculines (*-am*, *-iēm*). The Baltic adjectival declensions are paired for gender on the basis of vowel height, yielding two major types for Lithuanian, a low-vowel type, in which *a*-stem masculines are paired with *ā*-stem feminines (*gēras* : *gerà* ‘good’) and *ia*-stem masculines with *īā*-stem feminines (*žālias* : *žalià* ‘green’); and a high-vowel type, in which *u*-stem masculines are paired with *i*-stem feminines (*saldūs* : *saldì* ‘sweet’). The high-vowel type is distinctive only in certain case forms; otherwise, it merges with the *ja/iā*-type. It is quite productive in modern Lithuanian and includes borrowed words: *abstraktūs*, *-i* ‘abstract’. Latvian has only the low-vowel declensional type: *a*-stem *mazs* ‘small, m.’ (Lith. *māžas*) paired with *ā*-stem *maza* ‘small, f.’ (Lith. *mažā*), and *ia*-stem *zaļš* ‘green’ (Lith. *žālias*) paired with *īā*-stem *zaļa* (Lith. *žalià*). The *u/i*-stem type of Lithuanian (where it has already merged with the *ja/iā*-type in most case forms) has been replaced in Latvian by the *ja/iā*-stem type: Lith. *gilūs*, *gili* ‘deep’, but Latv. *dziļš* (\**dziļšas*), *dziļa*; Lith. *platūs*, *plati* ‘wide’, but Latv. *plašs* (\**platīšas*), *plaša*; Lith. *saldūs*, *saldì* ‘sweet’, but Latv. *saļds*, *saļda* (with generalized hard stem).

TABLE 11.8 DECLENSION OF *A/Ā*-STEM AND *U/I*-STEM SIMPLE ADJECTIVES IN LITHUANIAN

<i>a/ā</i> -stem: <i>māžas</i> , à ‘small’ (a.p. 4)				
	<i>singular</i>		<i>plural</i>	
	masculine	feminine	masculine	feminine
nom.	<i>māžas</i>	<i>mažà</i>	<i>maži</i>	<i>māžos</i>
gen.	<i>māžo</i>	<i>mažòs</i>	<i>mažų</i>	<i>mažų</i>
dat.	<i>mažám</i> , OLith. <i>mažámui</i>	<i>māžai</i>	<i>mažiems</i> , OLith. <i>mažiemus</i>	<i>mažóms</i>
acc.	<i>māžą</i>	<i>māžą</i>	<i>mažūs</i>	<i>mažás</i>
instr.	<i>mažù</i>	<i>mažà</i>	<i>mažaĩs</i>	<i>mažomĩs</i>
loc.	<i>mažamè</i>	<i>mažojè</i>	<i>mažuosè</i>	<i>mažosè</i>
<i>u/i</i> -stem: <i>saldūs</i> , -i ‘sweet’ (a.p. 3)				
	<i>singular</i>		<i>plural</i>	
	masculine	feminine	masculine	feminine
nom.	<i>saldūs</i>	<i>saldì</i>	<i>saldūs</i>	<i>saldžios</i>
gen.	<i>saldaũs</i>	<i>saldžiõs</i>	<i>saldžių</i>	<i>saldžių</i>
dat.	<i>saldžiám</i>	<i>saldžiai</i>	<i>saldiems</i>	<i>saldžióms</i>
acc.	<i>saldų</i>	<i>saldžią</i>	<i>saldžius</i>	<i>saldžias</i>
instr.	<i>saldžiu</i>	<i>saldžia</i>	<i>saldžiaĩs</i>	<i>saldžiomĩs</i>
loc.	<i>saldžiamè</i>	<i>saldžiojè</i>	<i>saldžiuosè</i>	<i>saldžiosè</i>

*Declension types for compound (definite) adjectives*

The category of definiteness is marked in adjectives by the historical affixation of the third-person personal pronoun (in the appropriate case) to the simple (indefinite) form: Lith. indefinite *naũjas*, *naujà* ‘new’ : definite *naujàsis* (\**naũjas*+*jis*), *naujóji* (\**naujã*+*jĩ*); indefinite *saldūs*, *saldì* ‘sweet’ : definite *saldùsis* (\**saldūs*+*jis*), *saldžióji* (\**saldjã*+*jĩ*). The tendency to merge the Lithuanian *u/i*-declension with the *ja/jã*-type, noted above, is extended still further in the definite adjective: thus nom. sg. f. *saldžióji* has replaced the reflex of \**saldĩ*+*jĩ* found in OLith. *saldĩi* (read *saldýji*). The corresponding Latvian forms are less transparent: *mazaĩs* ‘small’, *mazã*. Old Prussian may show a comparable formation in *pirmois* (m.), *pirmoi* (f.) ‘first’. Old Lithuanian shows a peculiar adjectival formation in which pronominal *-jis* is added to a case form of a noun: *danguie-iis ukinikas* ‘the heavenly farmer’, where *danguie* (*dangujè*) is the locative (inessive) of *dangùs* ‘heaven, sky’ (Schmalstieg 1988: 303).

Lithuanian has generalized the mobile accentual paradigm for disyllabic adjectives; remnants of old initial stress are found in the nominative singular of a number of *u*-stem adjectives with an acute base: *aiškus*, *aiški* ‘clear’, *lygus*, *lygi* ‘flat, even’, but the definite forms *aiškùsis*, *aiškióji*; *lygùsis*, *lygióji* follow the productive (mobile) pattern. Latvian did not generalize mobility in adjectives, and therefore preserves the original accentual distribution of acute bases.

TABLE 11.9 DECLENSION OF *A/Ā*-STEM AND *U/I*-STEM DEFINITE ADJECTIVES IN LITHUANIAN

<i>a/ā</i> -stem: <i>māžas</i> , <i>à</i> ‘small’ (a.p. 4)				
	<i>singular</i>		<i>plural</i>	
	masculine	feminine	masculine	feminine
nom.	<i>mažasis</i>	<i>mažoji</i>	<i>mažieji</i>	<i>māžosios</i>
gen.	<i>māžojo</i>	<i>mažosios</i>	<i>mažųjų</i>	<i>mažųjų</i>
dat.	<i>mažajam</i>	<i>māžajai</i>	<i>mažiesiems</i>	<i>mažosioms</i>
acc.	<i>māžąjį</i>	<i>māžąją</i>	<i>mažiuosius</i>	<i>mažąsias</i>
instr.	<i>mažuoju</i>	<i>mažąja</i>	<i>mažaisias</i>	<i>mažosiomis</i>
loc.	<i>mažajame</i>	<i>mažojeje</i>	<i>mažuosiuose</i>	<i>mažosiose</i>

<i>u/i</i> -stem: <i>saldūs</i> , <i>-i</i> ‘sweet’ (a.p. 3)				
	<i>singular</i>		<i>plural</i>	
	masculine	feminine	masculine	feminine
nom.	<i>saldūsis</i>	<i>saldžioji</i>	<i>saldieji</i>	<i>saldžiosios</i>
gen.	<i>saldžiojo</i>	<i>saldžiūsių</i>	<i>saldžiųjų</i>	<i>saldžiųjų</i>
dat.	<i>saldžiam</i>	<i>saldžiai</i>	<i>saldiesiems</i>	<i>saldžiūsioms</i>
acc.	<i>saldųjį</i>	<i>saldžiąją</i>	<i>saldžiuosius</i>	<i>saldžiąsias</i>
instr.	<i>saldžiuoju</i>	<i>saldžiąja</i>	<i>saldžiaisiais</i>	<i>saldžiūsiomis</i>
loc.	<i>saldžiūjame</i>	<i>saldžiojeje</i>	<i>saldžiūsiuose</i>	<i>saldžiūsiuose</i>

## Numerals

### *Cardinal and ordinal numbers*

Also agreeing in gender, number, and case with a nominal head, and therefore functioning as adjectives, are the numbers from one to nine and their compounds. Lith. *vienas*, *-à* ‘one’ and Latv. *viēns*, *-a* continue an East Baltic base \*vein-, with an unexplained *v-* in comparison with SCR. *in* ‘someone, other’ (which also confirms the East Baltic acute); OPruss. *ains*, *ainā*, *ainan* apparently continues an *o*-grade \*oyn-, which may also underlie the East Baltic forms. The corresponding ordinal is Lith. *pirmas*, *à*, Latv. *pirmais* (dial. *pīrmais*), OPruss. *pirmois*. Lith. *dù* (m.), *dvi* (f.) ‘two’ continues a dual form (m.) \*d(w)o-h<sub>1</sub>, (f.) \*d(w)o-ih<sub>1</sub>; Latv. *divi* (m., from \*duvi by regular sound change), *divas* (f.) adopts the masculine/feminine plural endings found in ‘four’ through ‘nine’; OPruss. *dwai* ‘two’ is also a plural form. The ordinal is Lith. *añtras*, *à*, Latv. *ùotrais*, OPruss. *antars*, *antrā*, continuing an Indo-European form for ‘other’. Lith. *trys*, Latv. *trīs* ‘three’ represent an *i*-stem plural formation \*triies < \*treies, which does not distinguish gender. For the ordinal, East Baltic agrees with Slav. (OCS) *treții* in having refashioned the stem as \*tret-: Lith. *trėčias* (\*trečias), *trečią*, Latv. *trešais*, *trešā*; the distinct OPruss. *tirts*, *tirtis*, nom. sg. f. *tirti* may represent an earlier \*třtyo-, or a remaking of a BSl. \*trit(i)jos.

Regular plural declensions on an adjectival pattern begin with Lith. *keturi*, *kėturios* ‘four’, Latv. *četri*, *četras* (in Indo-European terms, \*k<sup>w</sup>etur-) ‘four’; Latv. *četri* is

a replacement for earlier *cetri* (cf. the ordinal *četuřtāis*) under Slavic influence. Like Slav. (OCS) *četvьrtь*, the Lithuanian and Old Prussian ordinals show a different syllabification of the *-ur-* sequence: Lith. *ketviřtas*, OPruss. *kettwirts*: PIE \*k<sup>w</sup>etw<sub>ɹ</sub>-t-. Lith. *penkì, peñkios* ‘five’, Latv. *pieci, piecas* have as their base PIE \*penk<sup>w</sup>e; Lith. *šeši, šėšios* ‘six’, Latv. *seši, sešas* continue EBalt. \*seš- (Latvian through \*sesj-). A reduced-grade is found in the OPruss. ordinal (*w*)*usztis*, f. *uschtai*, Lith. dial. *ušios* ‘six-week lying-in period (after birth)’. These point to an unusual pre-Baltic alternation \*sweks- ~ \*uks-. Lith. *septynì, septýnios* ‘seven’, Latv. *septiņi, septiņas* represent a Baltic \*septin < PIE \*septm̥ (the long *-ī-* in the Lithuanian forms is secondary). The ordinals OPruss. *septmas*, OLith. *sekmas* (\*sep(t)mas; modern *septiñtas*) are continuations of PIE \*septmo-. The acuted *-ō-* of Lith. *aštuoni, aštuonios* ‘eight’, Latv. *astuōņi, astuōņas* reflects the final laryngeal of PIE \*h<sub>3</sub>ekteh<sub>3</sub>; the related ordinal, OLith. *ašmas*, OPruss. *asmus* (Modern Lith. *aštuñtas*, Latv. *astuōtaiš*), rebuilt on the pattern of ‘seventh’, directly corresponds to Slav. (OCS) *osmь*. Lith. *devynì, devýnios* ‘nine’, Latv. *deviņi, deviņas* (EBalt. \*devin, after \*septin ‘seven’) agree in their initial *d-* with Slav. (OCS) *devětь*,<sup>8</sup> the Old Prussian ordinal *newīnts* ‘ninth’ preserves the original nasal of PIE \*h<sub>1</sub>néwn.

Unlike ‘one’ through ‘nine’, Baltic ‘ten’ and its multiples, including the teens, function as nouns, governing (in general) the genitive case (unless the entire noun phrase is governed by another case). Lith. *dešimtis* and indeclinable *dėšimt* ‘ten’, Latv. *desmit* (dial. *desimt*), and OPruss. *dessimpts* continue a Proto-Baltic consonant stem \*dešimt (the base of which is PIE \*dek<sub>m̥</sub>), which, like the majority of such stems, joined the *i*-stems; the Lith. gen. pl. *dešimtų* and OLith. (Daukša) gen. sg. *deszimtes* are vestiges of the consonantal declension. Lith. *šimtas* ‘hundred’ and Latv. *simts* are direct continuations (except for the loss of neuter gender) of PIE \*(d)k<sub>ɹ</sub>mtóm. Lith. *tūkstantis* (m. *ja*-stem, dial. and OLith. f. *i*-stem) ‘thousand’, Latv. *tūkstuotis* (m. *ja*-stem) are related in an unclear way to OPruss. (acc. pl.) *tusimtons*, Slav. (OCS) *tysęšti, tysōšti* and Gmc (Goth.) *þūsundi*; the *-k-* of the East Baltic forms is secondary.

‘Eleven’ through ‘nineteen’, which (like ‘ten’, etc.) have noun syntax, are formed according to different models in Lithuanian and Latvian. Latvian follows a Slavic-like model of ‘one-after-ten’, ‘two-after-ten’ (*-pa-d(e)smit*, cf. OCS (*дѣна*) *na desęte*): Lith. *viėnpadsmìt* ‘eleven’ (‘one-after-ten’), *divpadsmìt* ‘twelve’ (‘two-after-ten’), *trispadsmìt* ‘thirteen’ (‘three-after-ten’). Lithuanian reflects, with its formant *-lika*, a “Germanic” model of “one left,” “two left,” seen in Goth. *ainlif* ‘eleven’, *twalif* ‘twelve’: Lith. *vienùolika* ‘eleven’ (‘one-left’), *dvýlika* ‘twelve’ (‘two-left’), *trýlika* ‘thirteen’ (‘three-left’), *keturiólika* ‘fourteen’ (‘four left’), etc. Lith. *dvidešimt* ‘twenty’ replaces an old dual phrase *dvi dėšimti*. The remaining multiples of ten, through ‘ninety’, are formed with the accusative plural feminine of the number together with an indeclinable *dešimt*: *trisdešimt* ‘thirty’, *kėturiasdešimt* ‘forty’, *peñkiasdešimt* ‘fifty’, and so on.

## The pronoun

### Personal pronouns

(Standard) Lithuanian and Latvian differ in the initial vowel of the first-person singular pronoun: Lith. *aš*, dial. and OLith. *eš* ‘I’, Latv. *es*; Old Prussian agrees with Lithuanian in showing mainly *as*, with a few instances of *es*. The final consonant of the Balto-Slavic etymon \*ež- must have devoiced in Baltic before the operation of Winter’s Law, since the Baltic forms show a short vowel; compare Slav. (OCS) *azь*, with the reflex of lengthening. The oblique cases are formed on the stem \*man-

OPruss. \*men- (with variants). The second-person singular, Lith. *tù* (shortened by Leskien's Law), Latv. *tu*, Old Prussian *tū*, *toū*, points to an acuted \*tū (PIE \*tuH). In the first-person plural, there is a length difference between OPruss. *mes*, matching Lith. *mēs* (with secondary lengthening) and Latv. *mēs*, which shows a long acute vowel. The oblique cases retain an initial *n-* in Old Prussian: gen. pl. *noūson*, dat. pl. *noūmans* (comparable to Slav. (OCS) *nasъ*, *namъ*), as opposed to Lith. *mūsų*, *mūms*, Latv. *mūsu*, *mums*. The second-person plural shows agreement across Baltic: Lith. *jūs*, Latv. *jūs*, OPruss. *ioūs*, pointing to CBalt. \*iūs. Lithuanian preserves an old dual form in *mūdu*, *mūdvi* (dialectally also *vėdu*, *vėdvi*) 'we two', *jūdu*, *jūdvi* 'you two'. The third-person pronouns differ across the Baltic languages. Lith. (m.) *jis*, (f.) *ji*, with correlates in High Latvian, continues an Indo-European deictic stem \*i- (\*is, \*ih<sub>2</sub>) with initial *j-* generalized from a thematized oblique stem *ja-*. In a number of Lithuanian dialects, *ānas* ~ *anàs*, *-à*, 'yon' serves as the third-person pronoun, matching in origin and function Slav. (OCS) *onъ*, *ona*. The Latvian pair *viņš*, *viņa* is of uncertain origin; OPruss. *tāns*, *tannā* ~ *tennā* undoubtedly reflects a conflation of the demonstrative \*ta- and \*anas (cf. also Pol., Cz. *ten* 'that'). Alone among the Baltic languages, Old Prussian shows an enclitic *di-*: *prowela-din* 'betrayed him'. An attributive (emphatic) pronoun is represented in Lith. *pàts* (< *patis*), *patì*, Latv. *pats*, *paša* (*pati*) 'oneself';<sup>9</sup> the Old Prussian equivalent is *subs*, not found in East Baltic.

TABLE 11.10 DECLENSION OF THE LITHUANIAN THIRD-PERSON PERSONAL PRONOUN *JĪS, JÌ*

	masculine		feminine	
	singular	plural	singular	plural
nom.	<i>jis</i>	<i>jìē</i>	<i>jì</i>	<i>jōs</i>
gen.	<i>jō</i>	<i>jū</i>	<i>jōs</i>	<i>jū</i>
dat.	<i>jám</i>	<i>jìems</i>	<i>jái</i>	<i>jóms</i>
acc.	<i>jì</i>	<i>juōs</i>	<i>jā</i>	<i>jàs</i>
instr.	<i>juō</i>	<i>jaīs</i>	<i>jà</i>	<i>jomìs</i>
loc.	<i>jamè</i>	<i>juosè</i>	<i>jojè</i>	<i>josè</i>

### Clitic pronouns

A reflexive clitic particle, Lith. *-s(i)*, Latv. *-s*, is joined to a verb to express middle voice: Lith. *keliūos(i)* '(I) get up', Latv. *ceļuos* '(I) rise, get up'. In verb forms with a preverb or other prefix (*ne-*, *ie-*, *be-*), the particle follows Wackernagel's rule in Lithuanian (see below): *atsikeliu* '(I) get up', inf. *atsikélti*; *nesikeliu* '(I) do not get up', while in Latvian the particle behaves as in simplex forms: *pieceļuos* '(I) get up', inf. *piecēlties* 'to get up'. In origin, the reflexive particle is a clitic form of the third-person dative pronoun, found in Old Lithuanian texts in both dative and accusative functions. The old texts also preserve the pronominal clitics 1 sg. *mi* and 2 sg. *ti*, which often appear in second position: *Pa-mi-rōdik* 'show me' (Modern Lith. *parōdyk mán*), *pa-ti-rōdisiu* '(I) will show you' (modern Lith. *parōdysiu táu*). An example of a reflexive enclitic in Old Prussian is *maitātun-sin* 'to nourish oneself'.

### Demonstrative pronouns

The Baltic demonstrative pronouns reflect a system of person deixis: 'this' (near the speaker), 'that' (near the hearer), and 'yon' (beyond speaker and hearer). Lith. *šis*, *ši*



(PBalt. \*šī) ‘this’, Latv. *šis, šī*, and OPruss. (m.) *schis*, like Slav. (OCS) *sb, si*, continue an Indo-European deictic stem \*ki-; the *š-* of the Latvian and Old Prussian forms represents a spread of \*sja- from the oblique cases. Lith. *tàs, tà* (PBalt. \*tâ) ‘that’ and Latv. *tas, tā*, like Slav. (OCS) *tv, ta*, represent generalizations of the oblique stem \*to- of the Indo-European demonstrative pronoun \*so, \*seh<sub>2</sub>, \*tod. The Old Prussian set (m.) *stas*, (f.) *stai* (\*stājī?), (n.) *sta* ‘this, the’ apparently represents \*sita-, analogous to Lith. *šitas* ‘this’, a conflation of *šis* and *tas*; compare Lith. *štaiš*, abbreviated from *šitaiš* ‘here’. Lith. *anàs, -à* ‘yon’, cognate with Slav. (OCS) *onv, ona*, serves dialectally (eastern Lithuanian and Žemaitic) as the third-person pronoun.

### Interrogative pronouns

The interrogative pronouns in Baltic continue an Indo-European stem \*k<sup>w</sup>o-: OPruss. (m., f.) *kas*, (n.) *ka*. The East Baltic forms do not distinguish ±human or ±animate: Lith. *kàs* ‘who, what’, Latv. *kas*. The interrogative/relative Lith. *kuris, kuri* ‘which, that’, Latv. *kurš, kura*, apparently represents the addition of a pronominal *jis* to Lith. *kurš*, Latv. *kūr* ‘where’, used (as it is in dialects) as a relative pronoun. Lith. *kat(a)ràs, -à*, Latv. *katrs, katra* ‘which of two’ continues a PIE \*k<sup>w</sup>otero-. The correlative series *t-* (demonstrative) : *k-* (interrogative) : *ī-* (relative), found in many Indo-European languages, is represented in Baltic by such forms as Lith. *tòks (tòks), tokià* ‘such a’ : *kòks (kòks), kokià* ‘what (sort of)’ : *jòks (jòks), jokià* ‘no, not a single (originally, ‘any (sort of)’), parallel to Slav. (OCS) *takv, kavv, jakv*. Latvian presents a somewhat different formation in *tàds, tāda* ‘such a’; *kāds, kāda* ‘what (sort of)’.

## The verb

### Verbal inflection

While the Baltic noun presents an archaic appearance, the verb has been radically refashioned and reduced to three tenses: present, past, and future. The present continues the Indo-European present, with the following basic stem types: an athematic type (see below), a thematic type (Lith. 1 pres. sg. *lipù*, 3 sg. *lipa* : *lipti* ‘to climb’), and “semi-thematic” types (Lith. *i*-stem 1 pres. sg. *sédžiu*, 3 sg. *sédi* : *sédėti* ‘to sit’ and Lith. \**ā*-stem 1 pres. sg. *prašau*, 3 sg. *prāšo* : *prašyti* ‘to ask’).<sup>10</sup> The future, formed on the infinitive stem with a suffixed *-sī/s-* (Lith. *kėlsiu, kėlsi, kėls* ‘(I, you, he/she/they) will raise’, Latv. *celšu, celsi, cels*; OPruss. *postāsei* ‘(you) will become’), continues an Indo-European desiderative.<sup>11</sup> The various Indo-European preterit formations have been replaced in Baltic by a simple innovative past. The function and forms of the Indo-European middle voice have been replaced by both an ablaut form (reduced grade) of a verbal root, and a reflexive construction: Lith. (trans.) *kėlti* ‘to lift, raise’, (intr./mid.) *kilti* ‘to rise’, (reflexive) *kėltis* ‘to get up (raise oneself)’ (the base here is PIE \*kelH- ‘to rise (up), tower (over)’, LIV<sup>2</sup> 349). Non-finite verb forms include the infinitive (the dative/locative of a pre-Baltic verbal noun in *-ti-*): Lith. *-ti*, Latv. *-t*, OPruss. *-t, -twei*; cf. CS *-ti*), vestiges of a supine (the accusative of a pre-Baltic verbal noun in *-tu-*): Lith. *-tu*, Latv. dial. *-tu*, OPruss. *-tun, -ton* (serving as infinitive), cf. CS *-tv*; and various participles, which combine with forms of ‘to be’ to build a system of relative tenses. In addition to the indicative, a number of moods are distinguished, including an imperative, conditional, and optative (permissive); these are presented below.

### *Athematic presents*

Although the only productive stem type in East Baltic is now the thematic (continuing an Indo-European stem with *-e/o-* formant), athematic verbs experienced a period of late productivity in Lithuanian, expanding beyond their original scope. Athematic presents are well attested in texts from the sixteenth and seventeenth centuries (see Petit 2010: 211–212) and as relic forms in dialects. Old Lithuanian ‘to be’ and ‘to give’ (with regularized spellings) serve as examples: ‘to be’ 1 sg. *esmi* (\**es-mie*), 2 sg. *esi* (\**es-sie*), 3 sg. *est(i)*; 1 pl. *esme*, 2 pl. *este*; 1 du. *esva*, 2 du. *esta*; ‘to give’ 1 sg. *duomi*, 2 sg. *duosi*, 3 sg. *duost(i)* (= OPruss. *dāsti*); 1 pl. *duome*, 2 pl. *duoste*; 1 du. *duova*, 2 du. *duosta*. A trace of the original athematic ending can be found in Latv. 1 sg. *ēsmu*, Lith. dial. *esmù* (standard Lith. thematic *esù*) ‘(I) am’. Best represented among the Old Prussian athematic verbs is ‘to be’: 1 sg. *asmai* (also *asmu*), 2 sg. *assei* (with variant spellings), 3 sg. *ast*; 1 pl. *asmai*, 2 pl. *astai*. The Lithuanian first- and second-person endings *-mi* and *-si* are Leskien’s Law shortenings of \**-mie*, \**-sie* (cf. OLith. *essiegu*, with emphatic particle *-gu*), corresponding to the *-mai*, *-sai* of Old Prussian. These may represent a conflation of Indo-European \**-mi*, \**-si* and the (originally middle-voice) \**-ai* of the 1 sg. perf. (OCS *vědě*), as inherited perfects were absorbed by the athematics in Baltic.

### *The thematic verb type*

Thematic verbs use the same set of person markers for all three tenses. The first-person singular marker *in*, for example, Lith. *keliù* ‘(I) lift, raise’, (refl.) *keliùos(i)* ‘(I) get up’, Latv. *ceļu*, (refl.) *ceļuôs* ‘(I) rise, get up’ directly reflects a PIE \**-oH*. The marker of the second-person singular (Lith. *keli* ‘lift, raise’, *keliēs(i)* ‘get up’; Latv. *cel* (\**celi*), *celiēs* ‘rise, get up’) is an EBalt. \**-êi* of uncertain origin. The Old Prussian material, due to the circumstances of its transmission, is not clear; Stang (1966: 408) suggests that forms such as *gīwassi*, *gīwasi* ‘(you) live’ are comparable to Slav. (OCS) *živeši* and represent a conflation of the athematic second-person singular ending *-si* with the \**-êi* of East Baltic. The endings of the first and second persons plural, Lith. *kēliame*, *kēliate*, (refl.) *kēliamēs*, *kēliatēs*; Latv. *ceļam*, *ceļat*; (refl.) *ceļamiēs*, *ceļatiēs*, reflect EBalt. \**-mê*, \**-tê*, the acuted length of which is perhaps by analogy to the singular forms. A distinctively Baltic innovation is the loss of a number distinction (singular/dual/plural) in the third person; here thematic verbs show a bare stem in *-a* for all persons: Lith. *kēlia* ‘lift, raise’, *kēlias(i)* ‘get up’, Latv. *ceļ*, *ceļās* ‘rise, get up’; past tense: Lith. *kēlē*, *kēlēs(i)*; Latv. *cēla*, *cēlās*. Old Prussian occasionally distinguishes a singular for both thematic and athematic verbs, present and past, by means of a suffix *-ts* (cf. *tas* ‘that’): *astits* (\**esti-ts*) ‘is’.

### *The “semi-thematic” verb type*

In addition to the athematic and thematic stem types (the latter with a number of subtypes), Baltic also preserves stem types (termed “halbthematisch” by Stang (1966: 319)) in which an unsuffixed present tense stem is associated with an extended infinitive stem. In the first, a present stem with an *-i-* theme vowel (an analogous type in Slavic has a long *-ī-*) is paired with an infinitive containing an *-ê-* formant. These often have stative meaning: Lith. 1 sg. *sédziu* (inf. *sédėti* ‘to sit’), 2 sg. *sédi*, 3 sg. *sédi*; 1 pl. *sédime*, 2 pl. *sédite*; compare OCS *sěždō*, *sědišb* : *sěděti*. This type is also characterized by reduced-grade root vocalism: Lith. 3 pres. *mini*, inf. *minėti* ‘to mention, remember’, compare

OCS *тънѣти*, *тънѣти* ‘to think’. In another “semi-thematic” type, an *o*-grade present tense with an *-ā*- stem formant is paired with an infinitive in *-ī-*; these have iterative meaning: Lith. *gāno* : *ganýti* ‘to pasture (cattle)’ (originally ‘to drive (cattle)’), Latv. *gana*, *ganīt*; Lith. *prāšo* : *prašýti* ‘to ask’, Latv. *prasa* : *prasīt*, and correspond to a Slavic formation preserving a present tense marker *-ī-*: OCS *гонѣти* : *гонѣти* ‘to chase, persecute’, *просѣти* : *просѣти* ‘to ask’.

### *The Baltic preterit*

The Baltic preterit is marked by the addition of an *\*-ē̃* or *\*-ā̃* suffix to a verb stem, which generally agrees in ablaut grade with the infinitive stem; the thematic endings were added directly to this preterit formant. Since the third person has no ending, it can be considered the unmarked form. An example of a preterit with the *\*-ā̃* formant is Lith. *lįpti* ‘to climb’: Lith. 3 *lįpo* (*\*lip-ā̃*); 1 sg. *lįpaũ* represents *\*lipā + \*-ũo*; 2 sg. *lįpāĩ*, *\*lipā + \*-īe*; 1 pl. *lįpome*, *\*lipā + \*-mē̃*; and 2 pl. *lįpote*, *\*lipā + \*-tē̃*. The *\*-ē̃* formant is found in Lith. *kėlti* ‘to lift, raise’: Lith. 3 *kėlė* (*\*kêlē̃*); 1 sg. *kėliau* is then *\*kêlē̃ + \*-ũo* (with *-ėu* to *-jāu*); 2 sg. *kėlei*, *\*kêlē̃ + \*-īe*; 1 pl. *kėlėme*, *\*kêlē̃ + \*-mē̃*; and 2 pl. *kėlėte*, *\*kêlē̃ + \*-tē̃*. Latvian has generalized the *\*-ā̃* formant, but *\*-ē̃* is still found dialectally and in texts from the sixteenth and early seventeenth centuries. The earlier *\*-ē̃* is reflected in the timbre of an *e/ē* root vowel, which preserves a higher articulation before former *ē̃*: corresponding to Lith. *kėlė* ‘lifted, raised’, *vėdė* ‘led’, *nėšė* ‘brought’, Latvian has *cēla* (*\*cêlễ*), *veda*, *nesa*, with /e/ rather than /æ/; compare *sēda* ‘sat’, with /æ/, representing *\*sêd-ā̃*, Lith. *sėdo*. The limited Old Prussian material suggests that the *\*-ē̃* and *\*-ā̃* formants were found here as well: *\*-ē̃*: *ymmi[ts]* (*\*imē̃-ts*) ‘took’, *\*-ā̃*: *prowela[din]* ‘betrayed (him)’.

### *Baltic diathesis and ablaut*

The distribution of *\*-ā̃* and *\*-ē̃* in Lithuanian depends for the most part on stem type and root vocalism, but is ultimately connected with a system of diathesis that characterized an earlier stage of Baltic: *\*-ā̃* is associated with intransitive verbs and *\*-ē̃* with transitive.<sup>12</sup> One of the striking features of the Baltic verbal system is the use of ablaut to mark the relation of transitive/intransitive. In a once highly productive type of diathetic marking, a verbal base with transitive meaning came to be characterized by a *ja*-present (a subtype of the thematic present) with *e*-grade in the root and an *-ē̃* formant for the preterit; the *e*-grade was extended by analogy to the infinitive. A correlated base with intransitive (often inchoative) meaning was characterized by reduced grade of the root, with a nasal-infix *a*-present and an *ā̃*-preterite. The infixed present tense stem is always characterized by circumflex tone, regardless of the tone of the infinitive.<sup>13</sup> Thus, the *e*-grade transitive Lith. *kėlti* ‘to lift, raise’ is correlated with a reduced grade intransitive *kilti* (*\*kĩ̃-*) ‘to rise’. The present tense of the latter is marked by a nasal infix, and the tone becomes circumflex: *kỹla* (i.e., *kỹ̃la* < *\*ki-ñ-l-a*); the preterit shows the *ā̃*-formant: *kilo*. In an environment where a nasal infix is not possible, such as an obstruent stem with long root vocalism, an *-st-* suffix is added to the present tense stem: Lith. *tõlta* ‘to move away’: *tõlsta*, *tõlo*.

### *Mood*

The Old Prussian imperative 2 sg. *immais* ‘take’, 2 pl. *imaiti* (: *ĩmt* ‘to take’) continues an Indo-European optative in *-oi-*, also reflected in the Lithuanian permissive, formed with

the proclitic particle *te-*: *tesākau* ‘may he/she/they say’, *tedarbiē* (< \*-ai) ‘may he/she/they work’. A more common formation in Modern Lithuanian is *te-* plus the third-person present: *tesāko*, *tedirba*. Old Prussian is unique in forming a permissive with the particle *-sei* (*pareyse* ‘may X come’), which may be connected with the sigmatic future (Stang 1966: 442). The imperative is formed in Lithuanian by adding the particle *-k(ite)* to the infinitive stem: *eik*, *eikite* ‘go!’ (inf. *eiti* ‘to go’). The Latvian imperative is identical to the second-person present; the plural is formed by adding *-iet* to the singular: *saki* ‘say!’, *sakiet* (inf. *sacīt* ‘to say’). The conditional mood of East Baltic derives from a *-bi-* (‘be’) element added to the supine, most clearly seen in forms such as Lith. 1 pl. *būtumime* ‘we would be’, 2 pl. *būtumbite* ‘you would be’. Most of the present-day forms are unanalyzable (e.g., Lith. 1 sg. *būčiau* ‘(I) would be’, *būtu* ‘he/she/they would be’). In Latvian, a single form identical with the old supine has been generalized for all persons and numbers: *ceļtu* ‘X would raise’. Characteristic of both Lithuanian and Latvian is the use of a participial form – in Lithuanian an active participle in the nominative case, in Latvian an indeclinable form derived from the present active participle – to mark reported speech or indirectly perceived action (evidential mood): Latv. *Viņš dzīvojujot* (sp.: *dzīvojot*) *Rīgā* ‘He is said to live in Riga’, Lith. *Jis gyvėnąs Vilniuje* ‘He is said to live in Vilnius’. Peculiar to Latvian is the use of the debitive (expressing obligation), formed with the particle *jā-* and the third-person present, with the logical subject in the dative: *man jābrauc* (‘to-me must-go’) ‘I must go’; the debitive of ‘to be’ is formed with the infinitive: *man jābūt* ‘I have to be’.

### Participles

Lithuanian and Latvian present an elaborate system of participles and gerunds. Among the more frequent participles (which decline like adjectives) are the present active, continuing a PIE *-nt-* participle: Lith. *rašąs*, stem *rašant-*, inf. *rašyti* ‘to write’; OLith. (nom. sg. f. def.) *santūi* ‘being’; Latv. *rakstuošs*: *rakstīt*; OPruss. *dilants* ‘working’; past active, reflecting a PIE perfect participle in *-ues-*, CBalt. \*-uens ~ \*-us-: Lith. (*pa*)*rāšęs*, stem (*pa*)*rāšius-*; Latv. (*uz*)*rakstījis*, stem (*uz*)*rakstījuš-*; OPruss. *dāuns* ‘having given’; present passive: Lith. *rāšomas*; Latv. *rakstāms*; OPruss. *poklausīmanas* ‘being heard’; and past passive, reflecting a PIE verbal adjective in *-to-*<sup>14</sup> Lith. (*pa*)*rašytas*, Latv. (*uz*)*rakstīts*; OPruss. *dāts* ‘given’. The past active participle is used, together with a finite form of the verb ‘to be’, to form a system of perfect tenses: Lith. *aš esū* (*pa*)*rāšęs* (m.), (*pa*)*rāšiusi* (f.) ‘I have written’, *aš buvau* (*pa*)*rāšęs* (m.), (*pa*)*rāšiusi* (f.) ‘I had written’; Latv. *es esmu* (*uz*)*rakstījis* (m.), (*uz*)*rakstījuši* (f.), *es biju* (*uz*)*rakstījis* (m.), (*uz*)*rakstījuši* (f.). Passive constructions are formed with the copula and the corresponding passive participle: Lith. *knygà būvo rāšoma* ‘the book was being written’, *knygà būvo parašyta* ‘the book was/had been written’. Old Prussian has a passive construction using *wīrst* ‘becomes’: *laikūts wīrst* ‘is held’.

Both Lithuanian and Latvian preserve the reflex of an earlier dative absolute in gerund constructions, which serve as a dependent clause when the subject differs from that of the main clause. The subject of the gerund appears in the dative; the gerund itself is a frozen form of an active participle in the dative (still found in the oldest Lithuanian texts): Lith. *sáulei* (dat. sg.) *tėkant* (pres. gerund) ‘as the sun is/was rising’ (‘to-the-sun rising’), *gimus* (past gerund) *vaikui* (dat. sg.) ‘when the child had been born’; Latv. *saūlei* (dat. sg.) *lecuot* (pres. gerund) ‘at sunrise’, *ziemai* (dat. sg.) *tuvuojuoties* (pres. gerund, refl.) ‘as winter is/was approaching’. When the subjects of the main clause and dependent clause are identical (and the verbal action is simultaneous), a special active participle,

formed in Lithuanian with *-damas*, in Latvian with *-dams*, is used: Lith. *skaitýdamas*, Latv. *lasýdams* ‘(while) reading’.

### Aspect

Aspect in Baltic is not as fully grammaticalized as it is in Slavic. While imperfective simplex verbs may be perfectivized through the addition of a lexically empty prefix (Lith. *rašýti* ‘to write, imperf.’ : *parašýti* (perfect.), Latv. *rakstít* (imperf.) : *uzrakstít* (perfect.)), the aspectual meaning of verbs formed with a lexical prefix, for example, Lith. *aprašýti* ‘to describe’, Latv. *aprakstít*, is typically contextual. In Latvian, certain verbs with spatial prefixes and perfective meaning have acquired corresponding phrasal imperfectives consisting of the simplex verb and a spatial adverb: *iziet* (perfect.) / *iét árā* (imperf.) ‘to go out’, *nùoņemt* (perfect.) / *ņēmt nuóst* (imperf.) ‘to take off’. Perfective verbs used in the present tense have non-progressive meaning (like the simple aspect in English): Lith. *Sáulè pàteka* (perfect.) *rytuosè ir nusilèidžia* (perfect.) *vakaruosè*, Latv. *Saùle paceļas àustrumuós un nùoriét riētumuós* ‘The sun rises in the east and sets in the west’.

### Verbal derivation

Among the more widespread verb types with a derivational suffix in both present tense and infinitive/preterit forms is *\*-ē-ja* (inf. *\*-ē-ti*), forming deadjectival verbs: Lith. *sēnas* ‘old’ : *senėti* ‘to grow old’, 3 pres. *senėja*; Latv. *vecs* : *vecēt*, 3 pres. *vecē*, cf. Slav. (OCS) *starь* : *starěti*, 3 pres. *starějetь*; this suffix is also used to form causatives in Latvian: *aúdzēt* ‘to cultivate, grow (trans.)’, 3 pres. *aúdzē* : *aúgt* ‘to grow (intr.)’. Another causative type is found in *-in-a/ā-*: Lith. *auginti*, 3 pres. *augina* ‘to raise (children)’ : *áugti* ‘to grow’; Latv. *aúdzināt* ‘to bring up, nurture’, 3 pres. *aúdzina* : *aúgt* ‘to grow’, OPruss. *auginnons* ‘brought up, raised’, *(po)augints* ‘brought up’. A third widespread type consists of denominal verbs in *-áuja* (inf. *-áuti*), cognate with Slav. *-újo* : *-ova-*: Lith. *keliáuti* ‘to travel’, 3 pres. *keliáuja*, 3 past *keliávo* : *kēlias* ‘road’, OPruss. *grikaut* ‘to confess’ (‘acknowledge one’s sins’) : *grík-* ‘sin’.

## SYNTAX

### On case usage

The nominative case, in addition to its basic functions as a marker of the grammatical subject and predicate noun, also marks the second argument in Latvian debitive constructions, where the logical subject is in the dative: *Man* (dat. sg.) *jápērk màize* (nom. sg.) ‘I have to buy bread’. The nominative may mark the direct object of an infinitive in Baltic impersonal constructions: Lith. *Langaĩ* (nom. pl.) *reĩkia uždarýti* (‘to close’) ‘The windows need to be closed’ (Schmalstieg 1988: 145–147), Latv. (from older texts) *ziņgs* (nom. sg.) *vajadzēs mazgāt* ‘the horse will need to be bathed’ (Endzelin 1922: 409). Characteristic of East Baltic is the use of a preposed adnominal genitive, often in place of a relational adjective: *lietúvių* (gen. sg.) *kalbà* ‘Lithuanian language’ (‘of-Lithuanians language’), *latviešu valoda* ‘Latvian language’; Lith. *áukso* (gen. sg.) *žiedas*, Latv. *zēlta* (gen. sg.) *grēdzēns* ‘a gold ring’ (*áuksas/zēlts* ‘gold’), a usage that already appears in early translations of Polish texts into Lithuanian: *dangaüs* (heaven, gen. sg.) *karalýstè* ‘heavenly kingdom’ (Pol. *królestwo niebieskie*). The genitive also marks agent or cause

in passive voice, a function inherited from Proto-Indo-European (Schmalstieg 1988: 180): Lith. *vėjo* (gen. sg. agent) *nulaužtà bėržo* (adnominal gen. sg.) *šakà* ‘a birch-tree branch broken off by the wind’, Latv. *vēja nùlàužts bėrza zars*. In Lithuanian, a partitive genitive may be found in the role of subject or object: Lith. *Miškè yrà vilkų* (gen. pl.) ‘There are wolves in the woods’, *Atvažiàvo svėčių* (gen. pl.) ‘(Some) guests arrived’, *Radaũ klaidũ* (gen. pl.) ‘(I) found (some) mistakes’. Arguments in negative existential constructions are in the genitive: Lith. *Nėrà žuviėš* (gen. sg.) ‘There is no fish’, Latv. *Nav zivs*, as are direct objects in negative sentences: Lith. *Aš neturiũ nėĩ brólio* (gen. sg.), *nėĩ seseĩš* (gen. sg.) ‘I have neither a brother nor a sister’, OLith. *Ne turek kitu Diewu* (gen. pl.) *prieg manes* ‘Do not have other gods before me’ (Schmalstieg 1988: 173); Latvian uses the accusative here: *Es neėdu màizi* (acc. sg.) ‘I do not eat bread’. The object of the now-rare supine is in the genitive: Lith. *Sėnis parėjo pietũ* (gen. pl.) *válgytu* ‘The old man returned to eat dinner’ (*pietũs* ‘dinner, nom. pl.’), a usage that has been extended to infinitive constructions replacing the supine: *Sėnis parėjo pietũ válgyti*. The dative, alongside its basic function denoting indirect object, is found in impersonal constructions: Lith. *Mán* (dat. 1 sg.) *reikia pagalbòs* (gen. sg.), Latv. *Man* (dat. 1 sg.) *vajag palīdzību* (acc. sg.) ‘I need help’; Lith. *Mán* (dat. 1 sg.) *šálta* ‘I am cold’ (with a historically neuter adjective), Latv. *Man saĩst* (inf. *saĩt* ‘to freeze’). For ‘have’, Latvian uses a “mihi est” construction with the dative, while Lithuanian has “habeo” with accusative: Latv. *Man* (dat. 1 sg.) *ir grāmata* (nom. sg.) ‘I have a book’ (negated: *man nav grāmatas* (gen. sg.) ‘I do not have a book’); Lith. *Aš* (nom. 1 sg.) *turiũ knygą* (acc. sg.). The dative is also used for the logical subject in Latvian debitive constructions: *Man* (dat. 1 sg.) *jàpėrk grāmata* (nom. sg.) ‘I have to buy a book’. The basic function of the accusative is to mark the direct object of a verb (Lithuanian switches to a genitive object in a negative sentence). The accusative often marks temporal and spatial expressions; the Lithuanian and Latvian words for ‘today’ are adverbs derived from the accusative for ‘this day’: Lith. *šiaĩdien*, Latv. *šũodien*. In Lithuanian, the instrumental, in addition to denoting ‘by means of’ (Lith. *rašýti pieštukũ* ‘to write with a pencil’), marks a predicate complement with verbs denoting a change of state: *Jis tãpo žurnãlo* (gen. sg.) *mėtu* (gen. pl.) *žmogumi* (instr. sg.) ‘He became the magazine’s man of the year’; modern Latvian uses the nominative here. In Latvian, the role of instrumental has been assumed by the preposition *ar* ‘with’, used with the accusative form in the singular (*rakstĩt ar zĩmuli* ‘to write with a pencil’) and the dative form in the plural. The locative (inessive) is used without a preposition in both Lithuanian (*Vilniuje* ‘in Vilnius’) and Latvian (*Rĩgã* ‘in Riga’); the Latvian locative has also assumed illative meaning, represented in modern Lithuanian by *ĩ* + accusative: *Vĩnš ienãca istabã* (Lith. *Jis iėĩo i kambarĩ*) ‘He entered the room’.

### Basic word order<sup>15</sup>

The relative order of the major syntactic constituents of a sentence in the modern Baltic languages is determined by information structure – whether the constituents function as theme or rheme. In a neutral declarative sentence consisting of S and V, S is typically the theme and V the rheme, with resultant SV word order: Lith. *Kãimas tebemiegójo* ‘The village was still sleeping’, Latv. *Jãnis laca* ‘John is reading’; the V may be a nominal predicate, with the copula not overtly expressed in the present tense: Lith. *Jis studeĩntas*, Latv. *Vĩnš students* ‘He is a student’. In a themeless sentence, in which both constituents function as rheme, the unmarked word order is VS: Lith. *Išãĩšo rĩytas* ‘Morning dawned’, *Gyvėno senėlis ir senėlė* ‘There (once) lived an old man and old woman’, Latv. *Dzĩvuoja*

*tēvs* ‘There lived a father’. In sentences consisting of three basic constituents, the pragmatically neutral and statistically most prevalent word order is SVO: Lith. *Jis myli savo motiną*, Latv. *Viņš mīl savu māti* ‘He loves his mother’. In Lithuanian, although the object may be highlighted by inversion (thus SOV), SVO and SOV are often in free variation (Ambrazas et al. 1985: 676), with SOV more characteristic of dialects, folklore, and formulaic speech: *Pirmi gaidžiai vėlniā* [O] *baīdo* [V] ‘The first roosters scare away the devil’. SOV is also typical for sentences with object pronouns: *Āš jī* [O] *pamačiaū* ‘I saw him’, and is the unmarked order, alongside SVO, if the object is the rheme: *Dēdē lāpē* [O] *nušóvė* [V] ‘Uncle shot a fox’, in answer to the question *Kā nušóvė dēdė?* ‘What did uncle shoot?’ If S is the rheme, the most natural order is OVS: *Kās nušóvė lāpē?* ‘Who shot the fox?’ *Lāpē* [O] *nušóvė dēdē* [S] ‘Uncle shot the fox’. Other word order permutations are possible, accompanied by expressive intonation. For non-context-bound sentences, in which all three major constituents function as the rheme, S and V are commonly inverted (cf. VS for corresponding two-component sentences), thus both VSO: *Pàkviētė* [V] *žvirblis* [S] *visūs paukšteliūs* [O] ‘A sparrow invited all the birds’, and OVS: *Priešus* [O] *skýrė* [V] *ūpė* [S] ‘A river separated the foes’.

### Phrasal word order

Adpositions in the modern Baltic languages precede the governed noun (prepositions), in harmony with a now-neutral SVO basic word order. Lithuanian preserves a few postpositions (Ambrazas et al. 1985: 685): *jō dėkà* ‘thanks to him’, as does Latvian: *viņa dēļ* ‘because of him’. A vestige of head-final position is found in the East Baltic local cases, formed with postpositions (see above). Other dependents generally precede the head: the neutral position of an adjective is before the noun it modifies, and an adnominal genitive also precedes its head noun, in harmony with an earlier SOV basic sentence structure assumed for Indo-European, but at odds with the prevailing SVO structure of the modern Baltic languages: Lith. *tėvo* (gen. sg.) *žōdžiai*, Latv. *tēva* (gen. sg.) *vārdi* ‘father’s words’ (*tėvas/tēvs* ‘father’). Partitive genitives, on the other hand, are postpositional: Lith. *būtelis alaūs* ‘a bottle of beer’, Latv. *pudele alus* (contrast *alaūs būtelis, alus pudele* ‘beer bottle’).

### Wackernagel’s Law in Baltic

Vestiges of Wackernagel’s Law of clitic placement can be found in several formations in East Baltic (see especially Petit 2010: 260–307). These include *tnesis* with a reflexive particle between a preverb and verbal root: Lith. *at-si-kėlti* ‘to get up’, the Latvian of traditional folk songs (*daīņas*): *Iz-sa gauži raudājuo-s* (with reflexive doubly marked; the corresponding form in modern Latvian would be *iz-raudājuo-s*) ‘I wept bitterly’ and dialectally (Endzelin 1922: 703–708), and in Old Lithuanian with clitic pronouns or expressive particles: *pa-mi-sakyk* ‘tell me’, *pa-gu-žįstat* ‘you know (emphatic)’. In Old Lithuanian, we find as a relic the placement of the inherited relative pronominal \*yo- between a prefix and root in definite adjectives and participles: *pa-io-prasta*, modern *pà-prasto-jo* ‘ordinary, gen. sg.’, *pra-iis-pūļes*, modern *pra-puoļes-is* ‘having perished’; compare the Slavic superlative formant \*na-jь-, Russ. *na-*. In Lithuanian, the emphatic particle *-gi-* occurs as an enclitic with adverbs and conjunctions: *kuřgi* ‘where (emphatic)’, *taīgi* ‘so, therefore’, *ařgi* ‘emphatic question particle’, *iřgi* (*iř* ‘and’ + *gi*, lexicalized as ‘also’), and between a preposition and following noun, yielding extended forms such as *nuōg* (*nuō* ‘from’ + *gi*), *iņg* (*iņ* [modern *į*] ‘into, to’ + *gi*), *ižg* (*iš* ‘out of, from’ + *gi*), *priēg* (*priē* ‘at’ + *gi*), which were particularly widespread in Old Lithuanian texts.

## FURTHER READING

Two traditional handbooks of the historical and comparative phonology and morphology of the Baltic languages are Endzelīns 1971, useful for derivational suffixes, and Stang 1966, with a section on the historical prosody of the Baltic languages. The lengthy introduction to Arkadiev 2015 (1–109) presents an up-to-date overview of the field of Baltic linguistics. For a synchronic and diachronic sketch of Lithuanian accentuation, see Lehfeldt 2001: 103–129. A number of individual topics, such as the Baltic neuter, are treated in depth in Petit 2010. Kabelka 1982, Dini 2000, and Sabaliauskas 2002 emphasize cultural and historical issues connected with the development of the Baltic languages. The first etymological dictionary of Baltic, Derksen 2015, includes in its introduction an overview of the Baltic languages and their dialects, with special attention to Baltic prosodic phenomena in comparison with Slavic. The classic etymological dictionary of Lithuanian, with information on other Baltic languages, is Fraenkel 1962–65. Smoczyński 2016 emphasizes Baltic ablaut relations, with corresponding Slavic material. An etymological dictionary of Old Lithuanian (through 1700) is now available (Hock 2015). Karulis 1992 is a traditional etymological dictionary of Latvian; etymological and dialectal data can also be found in the classic dictionary of Latvian, Mülenbachs-Endzelīns (Mülenbachs 1923–32), available online. Mažiulis 1981 is an annotated edition of all the Old Prussian texts. Etymological dictionaries of Old Prussian include Mažiulis 1988–97 and (for verb forms) Smoczyński 2005. Schmalstieg 1974 offers a grammar of Old Prussian, and Schmalstieg 1976 a review of literature on Old Prussian language, history, and culture. Endzelin 1922 is an exhaustive descriptive and historical grammar of Latvian and its dialects. For an overview of the phonology and morphology of the Latvian dialects in comparison with the standard language, see Gāters 1977. Mathiassen 1996 and 1997 are concise introductory grammars of Latvian and Lithuanian. For a more extensive presentation of Latvian grammar, see Fennell and Gelsen 1980 and, with diachronic notes, Forssman 2001.

## NOTES

- 1 The classic study, based on the evidence of hydronyms, is Toporov and Trubačëv 1962.
- 2 The \*ē, \*ō stage is apparently preserved in western Žemaitic: *pēns* ‘milk’, *dōna* ‘bread’, compared with the *ie*, *uo* of standard Lithuanian *pienas* (Latv. *piēns*), *dūona*.
- 3 For certain stem classes of the Latvian verb, final shortening of long vowels and diphthongs and apocope of short vowels leads to homonymy between the first-person singular present and past tense forms on the one hand, and the second-person singular and third person in the present tense on the other hand. Thus, the Lithuanian distinction *áugu* ‘(I) grow’ : *áugau* ‘(I) grew’ (*áugti* ‘to grow’), *rišù* ‘(I) bind, tie’ : *rišau* ‘(I) bound, tied’ (*rišti* ‘to bind’) is lost in Latv. *aūgu*, *risu*, and the forms distinguished in Lith. *riši* ‘(you) bind, tie’ : *riša* ‘(he/she/they) bind, tie’ merge in Latvian as *ris*.
- 4 Saussure’s Law originally functioned as an automatic phonetic rule, operating on any instance of stressed short or circumflex followed by acute. However, with the loss of the conditioning environment (phonemic tone on unstressed syllables), together with the effects of Leskien’s Law, the process has become a morphologically conditioned rule. This is most obvious in derivation, where the effects of analogical leveling are already apparent in the earliest texts.
- 5 Latvian broken tone is described by Lehiste (1972: 313) as “a change in the mode of phonation approximately in the middle of the syllable nucleus [. . .] realized as laryngealization or as a glottal stop.”



- 6 The choice of \*š, \*ž for Proto-Baltic is motivated by their subsequent development as *h* in early Baltic borrowings into Baltic Finnic: Finn. *heinä* ‘hay’ < CBalt. \*šēina-, cf. Lith. *šiėnas*.
- 7 According to Kortlandt (1977: 325–326), the original glottalic feature of the plain voiced stops, together with the inherited laryngeal (merging the PIE laryngeals), was reinterpreted as a prosodic feature of the syllable, continuing to the present in Latvian and Žemaitic (Lithuanian) broken tone acute.
- 8 For a phonetic account of this change, see Hamp 2006.
- 9 A related noun is found in Lith. *patis* ‘husband’, *pati* ‘wife’, Latv. *pats*, *pati*, cf. OInd. *pātih* ‘master, husband’, Gr. *πόσις* ‘husband’; and Lith. *vižšpats* ‘lord’, OPruss. *waispattin* ‘wife, mistress of the house (acc. sg. f.)’. Except for the full grade of the *vies-/vais-* element, the Baltic forms are identical to OInd. *višpātih* ‘master of the house’; cf. also OLith. *viešpatni* ‘mistress of the house’ and OInd. *pānī* ‘mistress’. Old Lithuanian shows traces of a consonantal stem: nom. pl. *viešpates*, gen. pl. *viešpatų* (Fraenkel 1962–65, vol. 2: 1245).
- 10 Note that the latter type is not distinguished formally from an \*-ā preterit: 1 sg. *prašaũ*, 3 sg. *prāšo* (*prašyti* ‘to ask’) is present tense; *radaũ*, *rādo* (*rāsti* ‘to find’) is past. A present in \*-ā will typically have a preterit in \*-ē: 1 sg. *prašaiũ*, 3 sg. *prāšē*, and thus the distinction between present and preterit is maintained within the same verb.
- 11 The first-person future forms Lith. *dúosiu*, Latv. *duošu* ‘(I) will give’ are essentially the thematic equivalents of OInd. *dāsyāmi* ‘(I) will give’. A Slavic vestige of this formation is found in the RussCS future participle of *byti* ‘to be’, *byšęšteje*, *byšęšteje* (≈ Lith. fut. act. ptcp. *būsiant-* : *būti* ‘to be’) and OCz. *probyšúčný* ‘useful’.
- 12 The origin of the \*-ē and \*-ā formants remains unclear. Stang (1966: 388) compares the Baltic *ā-* and *ē-*preterits to Slavic imperfect formations: Lith. *vėdė* ‘led’ and OCS *veděaxъ* ‘(I) led’, Lith. *riņko* ‘gathered’ and OCS *žbdaaxъ* ‘(I) waited’; see also Petit 2010: 249–254 and Derksen 2011: 208.
- 13 This morphological mechanism accounts for a number of apparent exceptions to Winter’s Law in Baltic, as an expected acute syllable is replaced by a non-acute (circumflex or short). For example, the short root vowel of Lith. *dubius* ‘deep, hollow’, repeatedly presented in the literature as an exception to Winter’s Law, is in fact associated with the vocalism of intransitive *dubti* (*duĩba*, *dũbo*) ‘to become hollow, sunken’, an ablaut-based refashioning of the Winter’s Law acute found in the Latvian cognate *duĩbjš* ‘deep’ and the Lithuanian transitive *dũobti* (3 pres. *dũobia*) ‘to hollow out’ (see Young 2011).
- 14 Baltic lacks the *-(e)no-* formant found for the past passive participle in Slavic and Germanic, although Endzelĩns (1971: 97) notes that Latv. dial. *vedene* ‘daughter-in-law’ (to *vest*, 3 pres. *vėd* ‘to lead; take a wife’) and the like may contain a vestige of this formant; cf. OCS *vedenь*, past pass. ptcp. to *vesti* ‘to lead’.
- 15 The following account of Lithuanian word order, together with examples, is based on Ambrazas et al. 1985: 672–691.

## REFERENCES

- Ambrazas, V.; Valeckene, A.; Valjulite, Ė.; Garšva, K.; Girdenis, A.; Knjukšta, P.; Krinickajte, S.; Labutis, V.; Lajgonajte, A.; Oginskene, Ė.; Pikčilingis, Ju.; Ruže, A.; Sližene, N. & Ulvydas, K. 1985, *Grammatika litovskogo jazyka*, Mokslas, Vilnius
- Arkadiev, Peter; Holvoet, Axel & Wiemer, Björn (eds.) 2015, *Contemporary Approaches to Baltic Linguistics*, De Gruyter Mouton, Berlin–Boston

- Birnbaum, Henrik 1970, "Four approaches to Balto-Slavic," in: Rūķe-Draviņa, Velta (ed.), *Donum Balticum*, Almqvist & Wiksell, Stockholm: 69–76
- Derksen, Rick 1998, "The distribution of the Old Prussian tones," in: Smoczyński, Wojciech (ed.), *Colloquium Pruthenicum Secundum*, Universitas, Kraków: 45–51
- Derksen, Rick 2011, Review of *Untersuchungen zu den baltischen Sprachen* by Daniel Petit, *Baltic Linguistics* 2: 203–210
- Derksen, Rick 2015, *Etymological Dictionary of the Baltic Inherited Lexicon*, Brill, Leiden–Boston
- Dini, Pietro 2000, *Baltų kalbos: lyginamoji istorija*, translated from the Italian by Henrikas Zabulis, Mokslo ir enciklopedijų leidybos institutas, Vilnius [Dini, Pietro U. 2014, *Foundations of Baltic Languages*, English translation by Milda B. Richardson, Robert E. Richardson, Vilnius University, Vilnius]
- Eklblom, Richard 1933, *Die lettischen Akzentarten*, Almqvist & Wiksells, Uppsala
- Endzelin, Jan 1899, "Über den lettischen Silbenakzent," *Beiträge zur Kunde der indogermanischen Sprachen* 25: 259–274
- Endzelin, Jan 1922, *Lettische Grammatik*, A. Gulbis, Riga
- Endzelīns, Jānis 1971, *Jānis Endzelīns' Comparative Phonology and Morphology of the Baltic Languages*, translated from the Latvian by William R. Schmalstieg and Benjaminš Jēgers, Mouton, The Hague–Paris
- Fennell, Trevor G. & Gelsen, Henry 1980, *A Grammar of Modern Latvian* (3 vols.), Mouton, The Hague
- Forssman, Berthold 2001, *Lettische Grammatik*, J. H. Röhl, Dettelbach
- Fortunatov, Filipp Fedorovich 1880, "Über accent und länge in den baltischen sprachen," *Beiträge zur Kunde der indogermanischen Sprachen* 22: 153–189
- Fraenkel, Ernst 1962–65, *Litauisches etymologisches Wörterbuch* (2 vols.), Carl Winter, Heidelberg; Vandenhoeck & Ruprecht, Göttingen
- Gāters, Alfrēds 1977, *Die lettische Sprache und Ihre Dialekte*, Mouton, The Hague
- Girdenis, Aleksas 2003, *Teoriniai lietuvių fonologijos pagrindai*, Mokslo ir enciklopedijų leidybos institutas, Vilnius [English translation: Young, Steven (trans.) 2014, *Theoretical Foundations of Lithuanian Phonology*, Eugrimas UAB, Vilnius]
- Hamp, Eric P. 2006, "Unexpected laryngeals in initial clusters of Baltic (-Slavic)," in: Aronson, Howard I.; Dyer, Donald L.; Friedman, Victor A.; Hristova, Daniela S. & Sadock, Jerrold M. (eds.), *The Bill Question: Contributions to the Study of Linguistics and Languages in Honor of Bill J. Darden on the Occasion of his Sixty-sixth Birthday*, Slavica Publishers, Bloomington, Indiana: 117–119
- Hock, Wolfgang (ed.), 2015, *Altlitauisches etymologisches Wörterbuch* (3 vols.), Baar, Hamburg
- Kabelka, Jonas 1982, *Baltų filologijos įvadas*, Mokslas, Vilnius
- Kallio, Petri 2008, "On the 'Early Baltic' loanwords in Common Finnic," in: Lubotsky, Alexander et al. (eds.), *Evidence and Counter-Evidence: Essays in Honour of Frederik Kortlandt*, Rodopi, Amsterdam–New York: 265–277
- Karaliūnas, Simas 1987, *Baltų kalbų struktūrų bendrybės ir jų kilmė*, Mokslas, Vilnius
- Karulis, Konstantīns 1992, *Latviešu etimoloģijas vārdnīca* (2 vols.), Avots, Riga
- Kortlandt, Frederik 1977, "Historical laws of Baltic accentuation," *Baltistica* 13/2: 319–330
- Kortlandt, Frederik 1985, "Long vowels in Balto-Slavic," *Baltistica* 21/2: 112–124
- Lehfeldt, Werner 2001, *Einführung in die morphologische Konzeption der slavischen Akzentologie*, 2nd ed., Verlag Otto Sagner, Munich
- Lehiste, Ilse 1972, "Some observations concerning the third tone in Latvian," in: Valdman, Albert (ed.), *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*, Janua Linguarum, Series Maior 54, Mouton, The Hague: 309–315

- LIV<sup>2</sup> = *Lexikon der Indogermanischen Verben*, Rix, Helmut (ed.), Dr. Ludwig Reichert Verlag, Wiesbaden 2001<sup>2</sup>
- Mathiassen, Terje 1996, *A Short Grammar of Lithuanian*, Slavica, Columbus, Ohio
- Mathiassen, Terje 1997, *A Short Grammar of Latvian*, Slavica, Columbus, Ohio
- Mažiulis, Vytautas 1966 (vol. 1), 1981 (vol. 2), *Prūsų kalbos paminklai*, Mintis (1966), Mokslas (1981), Vilnius
- Mažiulis, Vytautas 1988–97, *Prūsų kalbos etimologijos žodynas* (4 vols.), Mokslas/ Mokslo ir enciklopedijų leidykla, Vilnius
- Mažiulis, Vytautas 2004, *Prūsų kalbos istorinė gramatika*, Vilniaus universiteto leidykla, Vilnius
- McCluskey, Stephen C., Schmalstieg, William R. & Zeps, Valdis 1975, “The Basel epigram: A new minor text in Old Prussian,” *General Linguistics* 15/3: 159–165
- Mülenbachs, Kārlis 1923–32, *Mūlenbacha Latviešu valodas vārdnīca*, ed. by Jānis Endzelīns, Izglītības ministrija, Riga [electronic version: <http://tezaurs.lv/mev/>]
- Nau, Nicole 2011, *A Short Grammar of Latgalian*, Lincom Europa, Munich
- Nesselmann, Georg H. F. 1845, *Die Sprache der alten Preussen: an ihren Ueberresten erläutert*, G. Reimer, Berlin
- Petit, Daniel 2010, *Untersuchungen zu den baltischen Sprachen*, Brill, Leiden
- Sabaliauskas, Algirdas 2002, *Mes Baltai*, Gimtasis žodis, Vilnius
- Saussure, Ferdinand de 1896, “Accentuation lituanienne,” *Indogermanische Forschungen* 6 (Anzeiger): 157–166
- Schmalstieg, William R. 1974, *An Old Prussian Grammar: The Phonology and Morphology of the Three Catechisms*, The Pennsylvania State University Press, University Park, Pennsylvania
- Schmalstieg, William R. 1976, *Studies in Old Prussian*, The Pennsylvania State University Press, University Park, Pennsylvania
- Schmalstieg, William R. 1988, *A Lithuanian Historical Syntax*, Slavica, Columbus, Ohio
- Smoczyński, Wojciech 2005, *Lexikon der altpreussischen Verben*, Innsbrucker Beiträge zur Sprachwissenschaft, Innsbruck
- Smoczyński, Wojciech 2016, *Lithuanian Etymological Dictionary*, Peter Lang, Frankfurt
- Stang, Christian S. 1966, *Vergleichende Grammatik der Baltischen Sprachen*, Universitetsforlaget, Oslo [cf. also: Stang, Christian S. 1975, *Ergänzungsband: Register, Addenda und Corrigenda zur vergleichenden Grammatik der Baltischen Sprachen*, Universitetsforlaget, Oslo]
- Toporov, V.N. 2006, “Baltijskie jazyki,” in: Toporov, V.N.; Zav’jalova, M.V.; Kibrik, A.A.; Rogova, N.V.; Andronov, A.V. & Korjakov, Ju.B. (eds.), *Jazyki mira: Baltijskie jazyki*, Academia, Moscow: 10–50
- Toporov, V.N. & Trubačëv, O.N. 1962, *Lingvističeskij analiz gidronimov verxnego Podneprov’ja*, Akademija Nauk SSSR, Institut slavianovedenija, Moscow
- Vaillant, André 1936, “Le problème des intonations balto-slaves,” *Bullétin de la Société de linguistique de Paris* 37: 109–115
- Vaillant, André 1950, *Grammaire comparée des langues slaves*, vol. 1, IAC, Lyon–Paris
- Young, Steven 1994, “Endzelin’s Law and Acute Tone in Latvian,” *Linguistica Baltica* 3: 101–108
- Young, Steven 2011, “Winter’s Law and Baltic ablaut,” in: Sukač, Roman (ed.), *From Present to Past and Back: Papers on Baltic and Slavic Accentology*, Peter Lang, Frankfurt: 237–244
- Zinkevičius, Zigmās 1966, *Lietuvių dialektologija*, Mintis, Vilnius
- Zinkevičius, Zigmās 2005, *Lietuvių tautos kilmė*, Mokslo ir enciklopedijų leidybos institutas, Vilnius

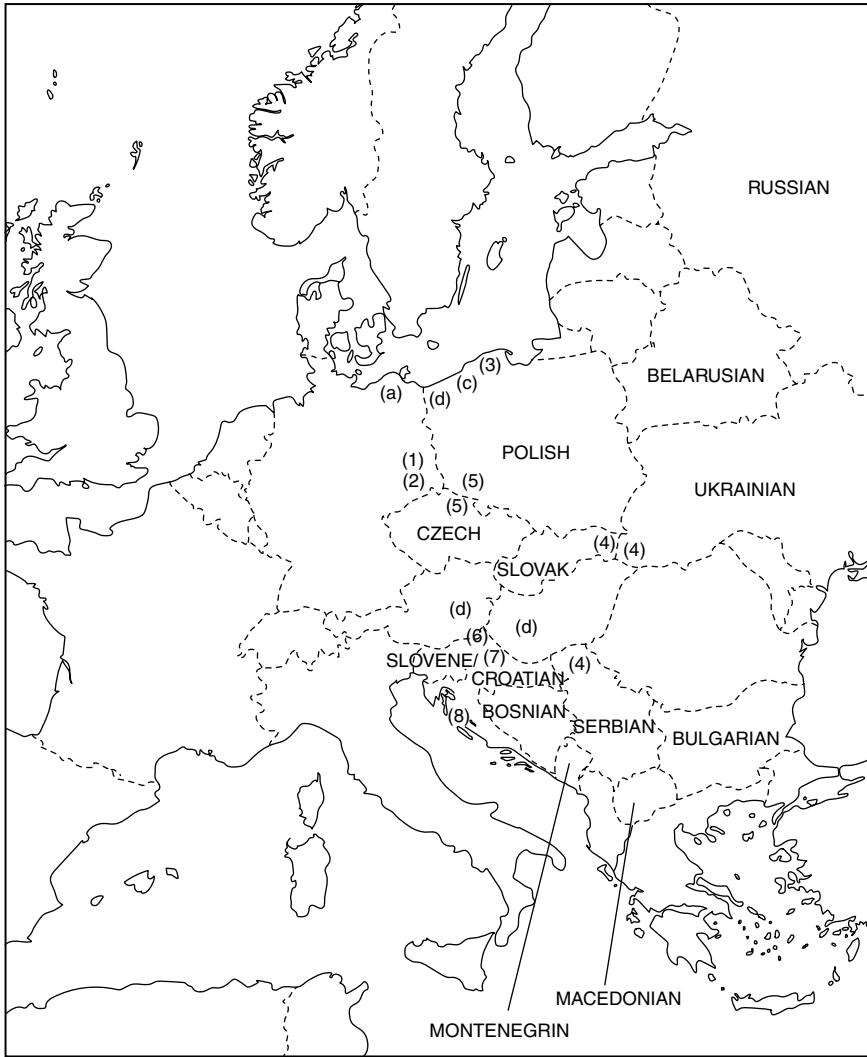
# SLAVIC

*Marc L. Greenberg*

## INTRODUCTION

The Slavic language family constitutes a widely distributed genetic grouping of languages spoken today from Central Europe to the Pacific, represented by standard languages that are traditionally divided into three branches: *West* – Czech, Slovak; Polish; Upper and Lower Sorbian; *South* – Slovene; Croatian, Bosnian, Serbian, Montenegrin; Macedonian, Bulgarian; and *East* – Belarusian, Ukrainian, Russian. The standard languages reflect ethnic and cultural distinctions that largely crystallized in the 19th century and as such are not ideal reflections of the genetic development of Slavic linguistic variation from the proto-language. For example, standard Croatian, Bosnian, Serbian, and Montenegrin are based on a single genetic dialect, Štokavian, but now are stylized in their officially valorized forms to fit the national identity projects of four post-Yugoslav states. Russian has a considerable lexical component of South Slavic origin, owing to its diglossic origins in the medieval period. Other Slavic language varieties – sometimes referred to as Slavic micro-languages – that have to a greater or lesser extent become codified but lack robust (or any) institutional support, exemplify further variegation, e.g., Kashubian (a West Slavic language spoken in the environs of Gdańsk), Rusyn (Carpathian variety of East Slavic with some West Slavic features), Silesian (transitional variety of West Slavic between the Czecho-Slovak sub-branch and Polish), Prekmurje Slovene (divergent variety of Slovene), and Kajkavian Croatian (non-standardized variety of Croatian with affinity to Slovene). Attested but extinct Slavic language varieties, for which descriptive information is available, include Polabian († mid-18 c.; westernmost Slavic language, along the Elbe/Laba river in today’s Germany), Pomeranian († medieval period; along the Baltic at the mouths of the Oder and Vistula rivers), and Slovincian († early 20 c.; northwestern Poland, where it was perhaps a variety of Kashubian). Another variety (or varieties) of Slavic existed in today’s Austria and Hungary, which, although unattested, is “Pannonian” Slavic.

Textual attestation of Slavic begins in the early second millennium A.D. with the Freising Folia (ca. 1000, written in Carolingian [Latin] script in a language with close affinity to what would later become Slovene) and the texts of the Cyrillo-Methodian tradition proceeding from the mid 19th c., though with extant manuscripts surviving from the 11th–12th cc. (Canonical “Old Church Slav[on]ic”). The latter are written in the Glagolitic alphabet, which was designed, based in part on existing alphabets, to reflect the sound structure of Slavic as it was spoken in the 9th c. around Thessaloniki; or its successor, the Cyrillic alphabet, which was based on contemporary Greek uncial script, mapping closely the (roughly) letter-to-phoneme correspondences of Glagolitic. An important and still growing body of texts are the Novgorod Birchbark Letters, which attest to the vernacular language (i.e., not influenced by the language of the church) of northern Russia



**MAP 11.2** MAP OF SLAVIC LANGUAGES

Titular/state languages are in ALL CAPS. *Other living varieties*: (1) Lower Sorbian, (2) Upper Sorbian, (3) Kashubian, (4) Rusyn, (5) Silesian, (6) Prekmurje Slovene, (7) Kajkavian Croatian, (8) Čakavian Croatian. *Extinct varieties*: (a) Polabian, (b) Pomeranian, (c) Slovincian, (d) Pannonian.

in the 12th–15th cc. and, by virtue of the dialect's peripheral position, illuminate archaic features of medieval Slavic. The language variation found in contemporary standard languages and dialects along with OCS serve as material for the comparative reconstruction of prehistoric stages of Slavic. In the exposition, focus will be on OCS forms, supplemented by material from living languages and dialects, where relevant.

## Periodization

Slavic is a *satem* language (cf. OCS *вѣсѣ* ‘village’, OInd. *viś-* ‘dwelling’, Alb. *vis* ‘locality’ || Lat. *vīcus*, Gr. οἶκος, Goth. *weihs*) and its closest congener the Baltic branch of Indo-European (see the chapter on Balto-Slavic). Periodization of Slavic is necessarily schematic before the period of the earliest attestations, which coincide with the beginning of disintegration into distinct dialect groups that can be more or less mapped onto the contemporary standard (“daughter”) languages. In this chapter we will operate with the following terminological designations:

*Proto-Balto-Slavic* – The emergent dialects whose innovations are distinct from Indo-European but are shared with dialects that were to become Baltic and Slavic.

*Proto-Slavic* – The emergent dialects whose innovations are distinct from Baltic but are shared with dialects that were to become Slavic. The reconstructions of PBSL and PSI are schematic and may be in some instances anachronistic because the time depth precludes precise dating. The representations in these systems are given in SMALL CAPS to make them visually distinct from CSI, which is the system underlying attested forms in living Slavic languages. Mutatis mutandis, *Proto-Baltic* refers to the Baltic forms paralleling PSI.

The term *Common Slavic* will be used in this chapter to refer to forms that underlie the earliest attested and contemporary Slavic languages as well as to the representations of these forms that are most familiar to scholars of Slavic, representing the vowel qualities presumed to be reflected in OCS (i.e., *i, y, u, ě, a; ę, q; e, o, ь, ъ*; see chart below).<sup>1</sup>

The Proto-Balto-Slavic period is contested in the literature with regard to the relationship between the two parts, i.e., whether Baltic and Slavic constituted a unified speech community, underwent common and parallel innovations because of proximity and/or bilingualism, or converged anew after a period of divergence. The facts cannot be interpreted unequivocally in favor of any one of these scenarios, so perhaps the most prudent, if unsatisfyingly equivocal, view is to consider the Proto-Balto-Slavic period one of a closer-knit interaction than with any other dialects emerging from Indo-European.

## Emergence of Slavic as a historical speech community, migrations

The Late Common Slavic period (ca. 500–1000 A.D.) coincides with the moment when a Slavic speech community appears relatively clearly in the historical record (mentioned by the Byzantines in 527; see Schenker 1995: 15ff.) as it expands from its core area and comes into contact with other speech communities. How the expansion proceeded has engendered considerable debate, ranging from a traditional view that sees the westward spread as owing to migration (e.g., Andersen 1996, 1999, 2003, Doluxanov 2000, Greenberg 2010) to a diffusion model that, in its fully articulated version, views the speech community as having developed a homogenized *lingua franca* used by both Avar elites and confederates, including the Slavs, in the Avar Khaganate (Lunt 1984–1985). For further discussion see Nichols 1993, Curta 2004, Snoj and Greenberg 2012.

On a micro-level the settlement pattern of Slavic communities centered on the *župa* (< \**gewpeh*<sub>2</sub>; Snoj 2003: s.v.) ‘a traditional district organized around a tribe’, which Pleterski describes as the building block of a “fractal society” of independent or loosely connected tribal units (2013: 10). These units were also the units of migration, which is reflected in the geographical dislocation and distal association of dialect features (see, for example, Andersen 1996, 1999, Schallert & Greenberg 2007).

In turn, the issue of whence the migrations originate has been the subject of controversy, though the clusters of Slavic hydronyms, in proximity with Baltic ones, found in the Dniester and Dnepr river systems of today's Ukraine provide a good anchoring point (Trubačev 1968), after which the Slavic speakers most likely spread by way of and across the Danube into central and southern Europe.

### Disintegration of Common Slavic and language contact

The second millennium A.D., often referred to as the time of “disintegration” of Common Slavic (Pol. *rozpad*, Russ. *raspad*), was a time of significant divergence among the Slavic languages owing to increasing geographical spread, linguistic drift, contact with Slavic and non-Slavic languages, and the rise and fall of local, confessional, and national identities. Structural effects that might be mentioned are the preservation of case marking in East Slavic and enrichment with additional case distinctions in Russian, brought about by contact with Finnic, viz. the partitive genitive and secondary (“concrete”) locative cases in standard Russian, which may be seen as just the tip of the iceberg with regard to convergent developments that have taken place among Baltic, Finnic, and northern Russian dialects. Central European Slavic languages, including the living West Slavic languages and the western South Slavic languages (especially Slovene and Croatian), have convergence features in morphosyntactic and lexical domains that are due to centuries-long contact with German, e.g., the use of prefixed imperfective verbs in narrated sequences of events (see Dickey 2011), or the development of a definite article, parallel to German *die/der/das* and Hungarian *a(z)*, in Czech (Doudleby dialect) *tí tahouni se používali po posekání obilí . . . a gdiš se strniště špatně posekalo, to tam bili tí kozi* ‘the draft horses were used after harvesting the crops . . . and when the stubble was poorly cut, then came the goats’ (Holub 2014: 183). The eastern South Slavic languages, Macedonian and Bulgarian, have undergone significant grammatical change as part of the Balkan convergence area. The dramatic nature of the structural change is heightened by the fact that in the earliest texts, OCS (in focus in this chapter) reflects the language state not long after it was relatively unified (“Common Slavic”), and the epicenter of Balkan features is in today's Macedonia, the home territory of the first Slavic language planners, Constantine and Methodius.

## PHONOLOGY

### Vowels

TABLE 11.11 PROTO-INDO-EUROPEAN TO LATE COMMON SLAVIC VOWELS

PIE	PSL.	LCSL.
	long vowels and diphthongs	/tense/ [long] vowels
*ī	*ī	<i>i</i>
*ey	*EY	
*oy	*AY	
*ay		
*ū	*ū	<i>y</i>
*ow	*AW	<i>u</i>
*aw		
*ew	*EW	<i>’u</i>

PIE	PSI.	LCSI.
	long vowels and diphthongs	/tense/ [long] vowels
*ē	*Ē	ě
*oy	*AY	
*ay		
*ō	*Ā	a
*ā		
*en, *em	*EN	e
*ṛ, *ṛ	*IM, *IN	
*on, *om	*AN	o
*ṛ, *ṛ	*UM, *UN	
*er	*ER	er
*or	*AR	or
*ar		
*el	*EL	el
*ol	*AL	ol
*al		
*r̥	*IR, *UR	r̥    ʀr, ʀr
*l̥	*IL, *UL	l̥    ʀl, ʀl
	short vowels	/lax/ [short] vowels
*i	*I	ɨ
*u	*U	ʊ
*e	*E	e
*o	*A	o
	Front	Back
High	ī/ı	ū/ʊ
Low	ē/e	ā/a

### *Monophthongization of diphthongs*

The transition from the PSI. to the CSI. stage, which underlies the attested and living systems of Slavic languages, is marked by the reshaping of syllables on the principle of “rising sonority” (see also below with regard to consonantal changes). In this set of developments, diphthongs monophthongized, and the functional load was transferred from quantitative to qualitative vowel contrasts. The equivalences are given in the chart above. The Late Common Slavic vowel system, which is close to that underlying OCS, was structured thus:

**TABLE 11.12 LATE COMMON SLAVIC VOWELS**

	front		back		
	tense	lax	tense	lax	tense
high		i	y		u
high-mid					ʉ
central		ь		ʊ	
mid		e		o	
low	ɛ	ě	a		



The monophthongization being a syllable-level phenomenon, diphthongs were treated differently in closed (tautosyllabic position) vs. open syllables (heterosyllabic), which gave rise to new alternations, e.g., \*plew-/plow- > PSI. \*<sub>PLAW-TĒY</sub>, \*<sub>PLA|U-E-XĪ</sub> > OCS *pluti*, *ploveši* ‘flow, swim’ inf., 2 sg.; PIE \*<sub>swe-kruh<sub>2</sub>-</sub> > PSI. \*<sub>SWEKRŪ</sub>, \*<sub>SWEKRU|W-E-S</sub> > OCS *svekyr*, *svekrŕve* ‘husband’s mother’ nom. sg., gen. sg. (cf. Sln. *kri*, *krvi*); PIE \*<sub>ken-/kŋ-</sub> > PSI. \*<sub>ZĀ-K(E)N-TĒI</sub>, \*<sub>ZĀ-KI|N-E-XĪ</sub> > OCS *začeti*, *začŕneši* ‘begin’ inf., 2 sg. With regard to relative chronology, the monophthongization followed the first palatalization of velars and preceded the second and third palatalizations. Moreover, the change of PIE \*<sub>ew</sub> > CSI. \*(j)u antedates deiotation following consonants, e.g., PIE \*<sub>lewd<sup>h</sup>eyes</sub> > PSI. \*<sub>LYAWD-IY-E</sub> > OCS *ljudije*; PIE \*<sub>kewmo-</sub> > PSI. \*<sub>SYAWM-A-S</sub> > OCS *šumъ* (also spelled *šjumъ*) (see also below). Note, however, that heterosyllabic \*<sub>EW</sub> merges with \*<sub>AW</sub>, e.g., PIE \*<sub>newo-</sub> > OCS *novъ* ‘new’.

### *Syllabic synharmony and palatalization*

The opposition of back-front was fundamental to the system, as is evident from the fact that (at first) consonants followed by front vowels were sub-phonemically palatalized. Palatalizations of velars and changes in clusters of *Cj* into palatal consonants (“deiotation”) introduced a systematic opposition between plain and palatal consonants (see further below). On the one hand, there was a tendency to adjust the tonality of vowels so that palatal consonants were followed by front vowels (“intrasyllabic harmony”), e.g., PIE \*<sub>moryo-</sub> > PSI. \*<sub>MAR-Y-A-(N)</sub> > OCS *morje*. The beginning of phonemic palatalization is perhaps attributable to the reinterpretation of PSI \*<sub>Ē</sub> (CSI. \*<sub>ě</sub>) as \*<sub>a</sub> after palatal consonants, which disrupted the pattern of intrasyllabic harmony, e.g., PIE \*<sub>g<sup>wh</sup>ēr-</sub> > PSI. \*<sub>GĒR-A-S</sub> > CSI. \*<sub>žarъ</sub> (Russ., Sln., BCMS *žar*). This points to the view that CSI. was sub-phonemically diphthongized [ʔæ], and subsequently raised to a high-mid diphthong throughout most of the Slavic territory, as is also indirectly recoverable from later reflexes of \*<sub>ě</sub>, cf. OCS *rěka*, Slk. *rieka*, Croat. *rijeka*. The onglide portion of the diphthong was evidently reinterpreted as a feature of the preceding consonant. Palatalization remained a fundamental element of the phonological system of Russian and Belarusian, whereas in Ukrainian and the West Slavic languages palatalization has been positionally eliminated and, in South Slavic in all but Bulg., eliminated altogether, save for the contrasts that arose with the loss of post-consonantal *j* (see below).

### *Elimination of diphthongs with r, l*

Sequences of vowel + *r*; *l* were eliminated in the LCSl. period and can be dated to the second half of the first millennium A.D., as evidenced by borrowings of toponyms assimilated during the migrations, e.g., Finn. *Alvatti* → Russ. *Lovat’* (river name), Finn. \*<sub>aal-dokas</sub> → Russ. *Ladoga* (lake name), Lat. *Albōna* → Croat. *Labin* (toponym), Lat. *Arba* → Croat. *Rab* (island name). The process ended by the 9th c., as borrowings related to Slavic Christianization show up without metathesis, e.g., Lat. *altāre* → OCS *olъtarъ* ‘altar’. Word-internal developments show dialect differentiation, where the open-syllable principle is realized differently for West and South Slavic, which show metathesis and lengthening of the vowel, and East Slavic, where the vowel is duplicated on either side of the consonant. The development was not carried through fully in Polabian. Examples: PIE \*(h<sub>2</sub>)<sub>wolko-</sub> → PSI. <sub>WALS-A-S</sub> > Pib. *vlās*, Pol., ULSorb. *wlos*, Cz., Slk. *vlas*, Sln. *las*, Mac., Bulg. *vlas*, OCS *vlasъ* || BRuss. *volas*, Ukr., Russ. *volos*; PIE \*<sub>g<sup>h</sup>ord<sup>h</sup>o-</sub> > PSI.

\*GARD-A-S > Plb. *gord*, Pol. *Białogard* (toponym in N. Poland) || Pol. *gród*, Cz. *hrad*, USorb. *hród*, Sln., BCMS, Mac., Bulg., *grad*, OCS *gradъ* || BRuss. *horad*, Ukr. *horod*, Russ. *gorod* ‘city’; PIE \*d<sub>er</sub>g<sup>h</sup>no- > \*PSI. DERN-A-S > Plb. *dren*, Cz. *dřín*, Slk. *drieň*, Sln. *dren*, Croat. *drijen*, Mac. *dren*, Bulg. *drjan* || BRuss. *dzěran*, Ukr., Russ. *deren* ‘cornus mas’.

### *Contraction, jer fall*

Among the final processes that marked the restructuring of Slavic are the loss of intervocalic *j* (“contraction”) and the “fall of the jers” (or “Havlík’s Law”), the alternating loss of the short lax vowels, *ь*, *ѣ*. These processes began ca. the 10th c. A.D. in West and South Slavic and progressed eastward, where contraction was limited in East Slavic and jer fall progressed well into the 12th c. Contraction resulted in increased frequency of long vowels in those areas where quantity distinctions remained phonemic, cf. CSI. \**moldaja* ‘young’ f. sg. def. > Cz., Slk. *mladá* || Sln., BCMS *mladā* || Russ., Ukr. *molodaja*; CSI. \**dělaješi* ‘you make, do’ > Cz. *děláš* || Sln. *delaš* || Russ. *delaěš* [d’elajiš].

Havlík’s Law identifies a process whereby word-final jers are lost universally and, by a process of compensatory lengthening, a jer in a syllable following a lost jer lengthens and merges with an existing vowel (with dialectally conditioned results). A jer in a syllable preceding a non-jer vowel elides. The change had wide-ranging effects on syllable structure, vowel systems, and word prosody, among the chief ones being the end of the centuries-long development of open syllables. Examples: CSI. \**къто* ‘who’ > Pol. *kto*, Croat. *tko*, Russ. *kto*, BRuss., Ukr. *xto*; CSI. \**въдовьсь*, \**въдовьса* ‘widower’ nom., acc./gen. sg. > Pol. *wdowiec*, *wdowca*, Cz. *vdovec*, *vdovce*, Sln. *vdovec*, *vdovca*, BCMS *udovac*, *udovca*, Mac. *vdovec*, Russ. *vdovec*, *vdovca*, Ukr. *vdovec’*, *vdovcja*; CSI. (with falling, “circumflex” stress) \**дѣнь* ‘day’ > Pol. *dzień*, Cz. *den*, Slk. *deň*, Sln., BCMS *dān* (long falling stress), Mac., Bulg. *den*, Russ., Ukr. *den’*, BRuss. *dzen’*; CSI. (with rising, “neo-acute” stress) \**рѣсь* ‘dog’ > Pol. *pies*, Cz., Slk. *pes*, Sln. *pes* [päs] (short falling stress), BCMS *pās*, Mac., Bulg. *pes*, Russ. *pēs*, BRuss. *p’os*, Ukr. *pes*.

### **Suprasegmentals**

The development of Slavic word prosody persists as one of the most dynamic areas of inquiry, and considerable debate surrounds its details. A basic account recognizes that PSI. developed a paradigmatic accent system with three patterns, which since Stang 1957 are referred to as accent paradigms (a.p.) A, B, and C. A.p. A is characterized by stable stem stress throughout the paradigm and is characterized by the “old acute” accent, which developed on long syllables typically of laryngeal origin. A.p. B has consistent stress on the ending alternating with “neo-acute” accent in the first syllable before the ending. The mobile a.p. C displays alternations between ending-final stress (or stress on the endings in acute-marked grammatical morphemes) in some forms and a falling tone (“circumflex”) on absolute initial syllables. In PSI. (and PBalt.) each syllable possessed a suprasegmental marking (or lacked it), which by LCSl. had resolved to a single stress for each phonological word.

A.p. A and B originate in PIE stem-stressed (barytone) paradigms and were differentiated by laryngeal (non-apophonic) length in stems of type A. After the application of Hirt’s Law, (“old acute”) stress was retracted or, in the case of two consecutive laryngealized syllables, assigned to the first of them, e.g., PIE \**k<sub>orh</sub>₂w-eh₂* > PSI. \**k<sub>ARW</sub>-A*

> CSI. *kǫrv-a* > Sln. *kráva*, Russ., Ukr. *koróva* ‘cow’, cf. PIE \*g<sup>w</sup>én-eh<sub>2</sub> > PSI. \*GEN-A > CSI. \*žen-ǎ > Russ. *žená* ‘wife, woman’. Non-acute stems underwent a forward shift of stress (Dybo’s Law), which was subsequently retracted in internal syllables (Stang’s Law), which in turn gave rise to the “neo-acute” rising-stress profile of a.p. C. The Slavic mobile type arose in PIE end-stressed and mobile paradigms (Meillet’s Law) with analogical loss of the laryngeal (“old-acute”) marking where applicable. The Slavic paradigms correspond to Lithuanian accent paradigms 1 (= A), 2 (= B), 3, 4 (= C) as in the following chart. Different scholars have treated the origin of the mobile pattern in Baltic and Slavic differently, with the two major schools – the Moscow School, led by Vladimir A. Dybo, and the Leiden School, led by Frederik Kortlandt – viewing the pattern as archaic (Moscow) vs. innovative (Leiden) with respect to Vedic and Greek. The exposition here cannot do justice to the complexities of the developments, for which reason the interested reader is referred to the critique in the appendix to Lehfeldt 2001 and to Kapović 2015 (esp. 195ff.) for the current state of the art, as well as the literature cited therein.

**TABLE 11.13 PIE, BALTIC, AND SLAVIC ACCENT PARADIGM CORRESPONDENCES**

Vedic, Greek, Germanic	Baltic	Slavic
barytone	1 – Lith. stem-stressed with acute tone; glottalized (“broken”) tone in Lith. Lowland (Žemaitian) dials. and Latvian	A – stem-stressed with “old-acute” stress; secondary traces of glottal stop or glottalization
	2 – Lith. stem-stressed non-acute tone with end stress on endings	B – ending-stressed alternating with neo-acute on stem-final syllable
oxytone	3 – Lith. mobile-stressed with “acute” tone in stem-stressed forms; glottalized tone in Lith. Lowland dials. and Latvian	C – mobile stressed, alternating between word-final stress and initial (recessive) “circumflex” stress
	4 – Lith. mobile-stressed with “circumflex” tone in stem-stressed forms	

“Old acute” stress is often marked with a double acute (ǎ) to distinguish it from the “neo-acute” (long: á, short: è) in order to avoid specifying the nature of the specific suprasegmental properties in question. The cover terms may be read as empty labels for suprasegmental contrasts. However, in this author’s view the laryngeal origin of the old-acute stress type can be read as “glottalized”; indirect evidence of its persistence in Slavic as either a glottalized (creaky-voiced) vowel or a glottal stop up through LCSl. is found in South Slavic pitch-accent systems: In Slovene in non-final stressed syllables, the tone contour is low in a phonetically long syllable (Sln. *bráta* ‘brother’ gen. sg.), while in final or monosyllabic words the tone is high in contrastively short syllables (Sln. *brát* ‘brother’ nom. sg., vs. long falling in “circumflex” *brát* ‘to go read’ sup.). In BCMS the reflex is short falling stress, e.g., *brät*, *bràta*. To wit, creaky-voiced syllables are correlated with low pitch and vowel lengthening, while glottal stops are correlated with high pitch and vowel shortening. In other words, on this view, which builds on Kortlandt’s view of the persistence of laryngeals in Baltic and Slavic (see Kortlandt 1985) creaky voice developed from the “old-acute”-marked syllables in non-final syllables in Sln., into

glottal stops in final or monosyllables and, in BCMS, in any old-acute-stressed syllable, and were subsequently reinterpreted upon the loss of the glottal feature as either long rising tone or short falling stressed syllables, respectively (Greenberg 2007, Holub & Greenberg 2013). In accord with tradition the sign for the long neo-acute (*á*) is used here, where it means a long rising tone, as is found, for example, in Kajkavian and Čakavian Croat. *súša* ‘drought’ (< LCSl. *súša* < PSI. \*SÁUX-Y-Ā). The grave accent (*è*) signifies the neo-acute in short syllables, e.g., LCSl. *kòlb* ‘stake’ (< PSI. \*KAL-ĀS), \*vòl’*a* ‘will’ (noun) (< PSI. \*VAL-Y-Ā), as well as stress in endings that is not of laryngeal origin, e.g., \*kolà ‘stake’ gen. sg. In the latter case, the signification recognizes that in examples where length has been preserved or extended to the stressed syllable, the surface forms in tonal dialects display rising pitch, e.g., Čakavian Croat. *noǵé* ‘of a foot’ gen. sg. In the case of the “circumflex,” the signs borrowed from Serbo-Croatian dialectology for the falling tone are typically used (long: *â*, short: *è*). Here, instead, for reconstructed Slavic forms the ictus mark (‘CV) is used to emphasize the default (phonemically unstressed) nature of the inherited Proto-Slavic circumflex, which is a fixed (falling-contour) stress in the first syllable of the phonological word, e.g., CSL. \*’nogǫ ‘foot’ acc. sg., \*’na nogǫ ‘onto (the) foot’. This configuration is preserved as such in, e.g., BCMS *nǒgu*, *nà nogu*, as well as reflected in the stress pattern of Russ. *nógu*, *ná nogu* (see Jakobson 1963: 71–72).

TABLE 11.14 ACCENTUAL PARADIGMS

Slavic a.p.	PIE	Baltic	Slavic
A	*bréh <sub>2</sub> tēr > Ved. <i>bhrátar-</i> , Gr. φρᾶτήρ ‘brother’ *d <sup>h</sup> uh <sub>2</sub> mó- > Ved. <i>dhūmá-</i> , Gr. θῆμῶς ‘smoke’ *seh <sub>1</sub> -ye/o- ‘sow’	Lith. <i>brólis</i> (a.p. 1), Latv. <i>brālis</i>    Lith. <i>sėti</i> , <i>sėjū</i>	CSl. *brātrь, gen. sg. *brātra > Sln. <i>brät</i> , gen. sg. <i>bráta</i> CSl. *dŷmь, *dŷma > Sln. <i>dīm</i> , <i>díma</i> CSl. *sějati, sějō, Croat. <i>sijati</i> , <i>sijēm</i>
B	*d <sup>h</sup> worom > Ved. <i>dvāram</i> ‘courtyard’  *wob <sup>h</sup> seh <sub>2</sub> > OHG <i>waspa</i> ‘wasp’ *g <sup>wh</sup> en-, Ved. <i>hānti</i> , Gr. θεῖνω ‘hit, drive (trans.)’	Lith. <i>dvāras</i> (a.p. 2)   Lith. <i>vapsà</i> (a.p. 2 > 4)  Lith. <i>giñti</i> , <i>genù</i>	CSl. *dvōrь, gen. sg. *dvōrā > Sln. <i>dvōr</i> , <i>dvōra</i> ; Croat. Čak. dial. <i>dvōr</i> , <i>dvōrā</i> CSl. *(v)osǎ, *(v)osǫ > Sln. <i>ósā</i> , <i>ósō</i> CSl. *gŷnāti, *ženǫ, *ženěty > Sln. <i>gnāti</i> , <i>ženem</i> , <i>žéne</i>
C	*suHnús, Ved. <i>sūnú-</i> ‘son’ *ōh <sub>2</sub> yóm, Gr. ὄιον ‘egg’ *ǵ <sup>h</sup> eym-eh <sub>2</sub> , Ved. <i>himā</i>	Lith. <i>sūnùs</i> (a.p. 3)   Lith. <i>žiemà</i> (a.p. 4)	CSl. *synь, Sln. <i>sīn</i> CSl. *’aje, Croat. dial. <i>jāje</i> CSl. *’zimǎ, *’zimǫ, Russ. <i>zimá</i> , <i>zīmu</i> , Sln. <i>zīma</i> nom. sg., <i>na</i> <i>zīmo</i> ‘in winter’ (acc.) CSl. *’mertī, *’mьrǫ, *’mьrěty > ORuss. <i>úmru</i> , <i>umrěty</i>
	*mer- ‘die’, Ved. <i>márate</i> ‘dies’	Lith. <i>miřti</i> , <i>miřtu</i>	

### Winter’s Law

In Baltic and Slavic an additional suprasegmental peculiarity, noted relatively recently (Winter 1978), is the appearance of acute-marked length in otherwise short-vowel syllables preceding a PIE voiced unaspirated stop, e.g., PIE \*sed- ‘sit’ > PBalt., PSI. *SĒD-Ē-TĒY* ‘sit’ inf. > Lith. *sėdėti*, OCS *sěděti* (vs. Lat. *sedēre*). Kortlandt (1978) has claimed that this feature corroborates Gamkrelidze and Ivanov’s reconstruction of the PIE consonant

system as consisting in PLAIN : GLOTTALIZED : ASPIRATED rather than the traditional PLAIN : VOICED : VOICED ASPIRATED (Gamkrelidze & Ivanov 1973; see also Vermeer 1984: 335, Kortlandt 1985), in which it is assumed that the glottalic feature for the consonant is transferred to the preceding vowel. Winter's Law is also correlated with the appearance of acute stress in the relevant syllable, suggesting that, if the glottalized interpretation of the old plain voiced series is correct, then the glottalization was transferred to the preceding vowel in Baltic and Slavic. In the chart below, the LCSL reflexes include later changes marked "cond." for conditioned changes, explained forthwith.

## Consonants

TABLE 11.15 PROTO-INDO-EUROPEAN TO LATE COMMON SLAVIC CONSONANTS

PIE	PSL.	LCSL.
*p	*P	<i>p</i>
*b	*B	<i>b</i>
*b <sup>h</sup>		
*t	*T	<i>t</i> , cond. <i>t'</i>
*d	*D	<i>d</i> , cond. <i>d'</i>
*d <sup>h</sup>		
*k	*K	<i>k</i> , cond. <i>č</i> , <i>ć</i>
*k <sup>w</sup>		
*g	*G	<i>g</i> , cond. <i>dž</i> > <i>ž</i> , <i>dž</i>
*g <sup>w</sup>		
*g <sup>h</sup>		
*g <sup>wh</sup>		
*ǵ	*Z	<i>z</i> , cond. <i>ž</i>
*ǵ <sup>h</sup>		
*k	*S	<i>s</i> , cond. <i>x</i> (> <i>š</i> , <i>ś</i> )
*S		
*m	*M,	<i>m</i>
*n	*N	<i>n</i> , cond. <i>ń</i>
*r, *l,	*R, *L	<i>r</i> , cond. <i>ř</i> ; <i>l</i> , cond. <i>l'</i>
*y, *w	*Y, *W	<i>j</i> , <i>v</i>

### *Elimination of geminates, dissimilation \*t/d > \*s before \*t*

Inherited geminate fricative clusters were eliminated, e.g., PIE \*skōy- → PSL. \*SĀI-NI > OCS *senb* 'shadow', PIE \*tuh<sub>2</sub>s-(d)kmt- > PSL. \*TŪSINT-Y-I > OCS *tysešti* 'thousand'. Clusters of two dentals stops become fricative + stop. This change is late PIE, as it is shared by Baltic as well as Indo-Aryan, Germanic, Celtic, Latin, and Albanian, e.g., PIE \*h<sub>1</sub>ed- → PSL. \*YĒS-TĒY > OCS *jasti* 'to eat'; PIE \*k<sup>w</sup>itti- > PSL. \*KIST-I > OCS *čbstb* 'honor'. A small number of examples point to the Slavic merger of PIE \*pt > PSL. \*TT > ST, e.g., ORuss./OCS *stryi* 'father's brother', ORuss. *Stribogb* 'Slavic pagan god name' (both derived from PIE \*ph<sub>2</sub>tr- 'father'); similarly ORuss. *nesterā* 'niece' (< PIE \*nept-), though PSL. \*TT > T is also observed, e.g., OCS *netii* 'nephew'.

### *RUKI change*

PIE \*s changes to x in PSL. following the segments R, U, K, I, a change that pre-dates the merger of the fricative reflex of PIE \*k in Slavic with PSL. \*s, e.g., PIE \*wřs- > PSL.

\*WIRX- > ORuss. *vbrxъ*, but PIE \*pork- PSI. \*PARSEN(T)- > OCS *prase* ‘pig’. The change parallels a similar development in Indo-Iranian (*ṣ*) and Baltic (*š*), as well as, possibly, Armenian (p. 206, 495, 438). Further examples: PIE \*snus-eh<sub>2</sub> > PSI. \*SNUXĀ > OCS *snъxa* ‘daughter-in-law’ (cf. OInd. *snuṣā*); PIE \*ksewbh- → PSI. XŪBĀ > Cz. *chyba* ‘mis-take’, Ukr. *xybáty* ‘to wobble’ (cf. OInd. *kṣūbhyati* ‘to shake, vacillate’); PIE \*MAYSA- > PSI. \*MAYX- > OCS *měxъ* ‘bladder, bellows’ (cf. Lith. *maišas* ‘bag’). The change fails to occur before an obstruent, giving rise to alternations as in OCS *rěxъ* ‘I said’ aor. vs. *rěsta* ‘they both said’ aor. < PSI. \*REK-S-AM, \*REK-S-TĀ. Word-initial *x-* developed from the generalization of the *x*-reflex from prefixed forms (as well as initial \*sk- > \*ks-), e.g., OCS (*pri*)*xoditi* ‘to walk’ (‘to approach’) ← PSI. \*PREY-SAD-Ī-TĒY (derived from PIE \*sed-/sod-/sd- ‘sit’).

### Loss of final consonants

Slavic underwent a series of changes often referred to as the “law of rising sonority” or the “law of open syllables,” which eliminated syllables culminating in a segment other than a vowel. The changes developed from the elimination of consonants in word-final position, likely in the first centuries A.D., to the elimination of liquid diphthongs in the last centuries of the first millennium (the type PSI. \*GARD-AS > OCS *gradъ* || ORuss. *gorodъ* ‘fortification, city’). These changes included the monophthongization of diphthongs and the development of prothetic consonants, as well as entailing structural changes in grammatical endings. These issues will be discussed below in the appropriate sections. Here we will focus on some of the early changes in Slavic. The earliest change is loss of final \*-r, which is shared with Baltic, e.g., PIE \*meh<sub>2</sub>tēr → PSI. \*MĀTĪ, > OCS *mati* ‘mother’, PBalt. \*MĀTĒ > Lith. *môtė* ‘wife’.

### Cluster simplification

In addition to the adjustments in clusters mentioned above, remaining clusters and those arising in word formation were simplified, the result of which is the elimination of all clusters aside from fricative + stop, cf. PIE \*yeh<sub>2</sub>- ‘travel’ + \*-sd- ‘sit’ (Ø-grade) → PSI. \*YĒZD-Ī-TĒY ‘to ride’ > OCS *jazditi* ‘travel’, BCMS *jězditi* ‘to ride a horse’ [Greenberg and Dickey 2006]), though syllable onsets could also end in a sonorant, e.g., PIE \*h<sub>2</sub>eyskreh<sub>2</sub> > PSI. YĪSKR-Ā > OCS *iskra* ‘spark’. Examples: CC > C̄C – PIE \*wob<sup>h</sup>hseh<sub>2</sub> > \*PSI. AS-Ā > OCS *osa* ‘wasp’; PIE \*h<sub>1</sub>rouwd<sup>h</sup>so- > PSI. \*RAWS- > OCS *rusъ* ‘red’; h<sub>1</sub>dsnt-sni- > PSI. \*DENSN-I > OCS \*dęsnъ ‘gums’; PIE \*lowksneh<sub>2</sub> > PSI. \*LAWN-Ā > OCS *luna* ‘moon’; \*d<sup>h</sup>ǵ<sup>h</sup>emi ‘earth’ loc. sg. → \*ZEM-Y-Ā > OCS *zemlja* ‘earth’; PIE \*d<sup>h</sup>olb<sup>h</sup>- → PSI. \*DAL-TA > OCS *dlato*, Croat. *dlijeto* ‘chisel’; CC > C̄C – PIE \*ob- + \*h<sub>2</sub>welk-/ h<sub>2</sub>wolk- ‘drag’ → PSI. \*AB-ALK- > OCS *oblakъ* ‘cloud’; PIE \*nok<sup>w</sup>ts → \*NAKT-I- > OCS *noštъ* ‘night’; PIE \*wronk-eh<sub>2</sub> > PSI. \*RANK-Ā > OCS \*rōka ‘hand’.

The cluster \*sr developed excrescent *-t-* in conformity with the cluster constraint adduced above, an innovation also found in Baltic, Albanian, Germanic, and Celtic, e.g., PIE \*srew-/srow- ‘flow’ → PSI. \*STRAW-G-Ā, \*STRAW-M-Ū, -M-EN- > OCS *struga* ‘flow’, *Strumień* ‘hydronym in Poland’, cf. Eng. *stream*. The constraint remained productive into the historical period, e.g., OCS *Izdrailъ* ‘Israel’; *vъzdrydati* ‘to break out crying’ perfect., cf. *rydati* ‘to cry’.

A particular case is -gt-/kt- before *ī/i* (and later PSI. *y*), which merged with reflexes of CSI. \*t̥, \*d<sup>h</sup>ugh<sub>2</sub>tēr- > PSI. \*DUKT-Ī > OCS *dъšti* ‘daughter’, Russ. *doč’*, OCz. *dci*, for which reason \*/kt/ was likely realized as \*[t̥] (see Vermeer 2014: 217 and below).

TABLE 11.16 CLUSTER SIMPLIFICATION

CC > C	CC > CC
PS, TS, KS > S (X)	VR-, VL- > R, L
TK, DG > K, G	BV > V
KT > T [t̥]	
PT > ST, T	
TM, TN, DM, DN > M, N	

*First palatalization of velars*

Among the first Slavic innovations exclusive of Baltic is the palatalization of velars \*κ, G, X before front vowels, which occurred after the RUKI change and before the monophthongization of diphthongs. This places the change in the first half of the first millennium A.D. Borrowings from Slavic into Greek show that it was completed by the 7th c. A.D. (e.g., *Čbrnica* > Τσερνίτσι ‘toponym in Messenia’, \*Stružija > Στρούζια ‘toponym in Phoeis’, \*Vbršьskъ > Βερσίτσι ‘toponym in Achaia’). The reflexes are uniform in Slavic: PSI. \*κ, G, X > CSI. \*č, ž, š. It is generally considered that \*ž had originally paralleled the development of \*č by having an (allophonic) affricate stage *dž*, which is observed only indirectly in clusters of \*zG before a front vowel, e.g., PSI. \*DROZG-IY-Ā > CSI. \*droždžьja > OCS *droždija* (with other vocalic affixes) Ukr. *droždzi*, Prekmurje Sln. *droždže* ‘yeast’. Accordingly, \*sk also resulted in a combination of fricative + affricate, e.g., PIE \*h<sub>2</sub>eys-ske- > PSI. \*AYSK-E-MAS > Sln. *iščemo* ‘seek’ 1 pl.

TABLE 11.17 FIRST VELAR PALATALIZATION

PIE	PSI.	OCS
*kelH-	*KEL-A	<i>čelo</i> ‘forehead’
*pekʷ-	*PEK-E-XI	<i>pečesi</i> ‘bake’ 2 sg.
*gʷh <sub>1</sub> Hsl-	*GIL-A	<i>žila</i> ‘vein, nerve’
*magʰ-	*MAG-E-XI	<i>možeši</i> ‘can’ 2 sg.
*kel- > Gmc *χelma-z →	*XELM-	<i>šlěmъ</i> ‘helmet’
*dʰwes- →	*DAWX-Y-Ā	<i>duša</i> ‘soul’

*Prothesis*

Prothetic consonants in initial syllables where the inherited onset was vocalic were inserted, though differing results occur in the daughter languages owing to later developments. Examples: PIE \*h<sub>1</sub>nh<sub>3</sub>men- > PSI. \*INMEN- > CSI. \*jbmę > OCS *imę*, OCz. *jmě* ‘name’; PIE \*h<sub>2</sub>eskreh<sub>2</sub> > PSI. \*ĀISKR-Ā > CSI. \*jiskra > OCS *iskra*, Cz. *jiskra* ‘spark’; PIE \*d̥nh<sub>3</sub>uh<sub>2</sub>- → PSI. \*ENZŪK-A-S > \*jęзыкъ > OCS *jęзыкъ* ‘tongue, language; heathen’; PIE \*n̥ ‘in’ > PSI. UN > CSI. \*vъ > OCS *vъ* ‘in’; PIE \*ūpso-ko- > PSI. \*ŪSAK-A-S > CSI. \*vysokъ > OCS *vysokъ* ‘high, tall’ nom. sg. m. indef.

*Second and third palatalization of velars*

If the monophthongization of diphthongs marks the beginning of Common Slavic, the second and third palatalization of velars and “deiotation” give rise to isoglosses cutting across the Slavic speech areas, marking the disintegration of Common Slavic. Traditionally, the development has been broken into two, presumably chronologically distinct, changes. However, the second and third palatalizations were likely a single process with

two sets of conditioning environments: the “second” palatalization is triggered by the rise of new front vowels arising after the monophthongization of diphthongs (see above) and the “third” by a preceding front vowel.

TABLE 11.18 SECOND AND THIRD VELAR PALATALIZATIONS

PIE	PSl.	Attestations	Conditioning environment
*k <sup>w</sup> oy <sup>n</sup> -eh <sub>2</sub>	*KAYN-Ā	OCS <i>cěna</i> ‘price’	*K before front V (“2 pal.”)
*k <sup>w</sup> oyto-	*KWAYT-A-S	OCS <i>cvětb</i> ‘flower’, ‘color’    Pol. <i>kwiat</i> , Cz. <i>květ</i>	
*w <sub>1</sub> k <sup>w</sup> o-	*WILK-AYXU	ORuss. <i>vьlcěxb</i> ‘wolves’ loc. pl.	
*g <sup>h</sup> oylo-	*GAYL-A	OCS <i>ǰělo</i> ‘very’, Sln. <i>zelo</i>	
*g <sup>h</sup> woy-sd <sup>-2</sup>	*GWAYZD-Ā	OCS <i>ǰvězda</i> ‘star’    Pol. <i>gwiazda</i> , Cz. <i>hvězda</i>	
*ksoy-ro-	*XAYR-A	OCS <i>sěrb</i> ‘blue-gray’    Pol. <i>szary</i> , Cz. <i>šerý</i>	
	*GRAYX-AY	OCS <i>grěsi</i>    Pol. <i>grzeszy</i> , OCz. <i>hřieši</i>	
*meh <sub>1</sub> s-	*MĒS(E)NK-A-	OCS <i>měšecb</i> ‘moon’	*K after front V (“3 pal.”)
*stigho-	*STIG-Ā	OCS <i>stbǰa</i> ‘path’	
*wiso-	*WIX-A	OCS <i>vьsb</i> ‘all’    Pol. <i>wszystek</i> , Cz. <i>všechn</i>	

The diatopic variation found in the palatalization reflexes reveals a dynamic period of change, where the differing rates of expansion of the innovations result in varying relative chronologies (see esp. Andersen 1969, Vermeer 2014). Alongside the second and third palatalizations, the lenition of CSL. \*g > γ was under way in much of central Slavic, ultimately affecting Cz., USorb., Slovak, NW Sln., part of Čakavian Croat., BRuss., Ukr. and southern Russ. These are classic examples of center-periphery phenomena. For example, in the Novgorod-Pskov area (peripheral), as attested in the Birchbark Letters, e.g., *kěle* ‘whole’ nom. sg. m., *xěri* ‘of a gray cloth’ gen. sg. f. (cf. modern Russ. *celyj*, *seryj*), indicates that the second palatalization failed to reach this area. In West Slavic (central) the results of the palatalization of \*k, g are identical to the results of the deiotation of \*t, d (see below). If it is assumed that the second and third palatalizations resulted in [tʰ, dʰ] (palatal stops) in WSl., then the change of \*x in the second and third palatalization to š follows the systemic phonemicization of the series as palatals (rather than alveo-palatals, as elsewhere) (Vermeer 2014: 189). This scenario is confirmed by the fact that in West Slavic only the reflexes of the second palatalization before *v* have been phonologized as *k*, *g* (Pol. *kwiat* ‘flower’, *gwiazda* ‘star’). Had they been affricated, as elsewhere, depalatalization as stops would have been unexpected.

#### *Loss of post-consonantal j (“deiotation”)*

The last major CSL. change was the simplification of clusters of C + *j*, which in all cases resulted in the elimination of *j* (whence the traditional term “deiotation”). Dental stops were first palatalized and then, in another wave of innovation, assibilated. The early phase is attested in the Freising Folia, where we find the spelling *imoki* (2x) (< CSL. \*jьmofъjь



'having' pres. act. ptcp. nom. m. sg.); cf. in the same text *v ueki* (*vъ věky* 'for all time' acc. pl.). Dental fricatives merged with palatal fricatives from the first palatalization of velars (wherever an affricated realization of \*dž < \*<sub>G</sub> + front vowel had not been retained). After labials, *j* became *l*, and the sonorants *n*, *l*, *r* developed palatal counterparts. Variation is significant, as can be seen in the following chart.

TABLE 11.19 'DEIOTATION' REFLEXES

	PIE	CSL.	West	South	East
*tj > *t'	*kwoyt-o- 'bright'	*svěť'a 'candle'	Pol. <i>świeca</i> , Cz. <i>svíce</i> , USorb. <i>swěca</i>	Sln. <i>sveča</i> , BCMS <i>sviječa</i> , Mac. <i>sveka</i> , Bulg. <i>svešta</i>	BRuss. <i>svjača</i> , Ukr., <i>sviča</i> , Russ. <i>sveča</i>
*dj > *d'	*h <sub>1</sub> ud <sup>h</sup> - 'red'	*ръд'а 'rust'	Pol. <i>rdza</i> , Cz. <i>rez</i> , Slk. <i>hrdza</i> , LSorb. <i>rza</i>	Sln. <i>rja</i> , BCMS <i>rda</i> , Mac. <i>rġa</i> , Bulg. <i>ръзда</i>	Russ. <i>ržavčina</i> , BRuss., Ukr. <i>rža</i>
*sj > *š	*h <sub>1</sub> nek-/ h <sub>1</sub> nok- 'carry, take'	*noša 'something carried, worn'	Pol. <i>nosze</i> , Cz. <i>nůše</i>	Sln. <i>noša</i>	Russ., BRuss. <i>noša</i>
*zj > *ž	*h <sub>2</sub> eg- 'goat'	*koža 'skin, leather'	Cz. <i>kůže</i> , Slk. <i>koža</i> , ULSorb. <i>koža</i>	Sln. <i>koža</i> , BCMS <i>koža</i> , Mac., Bulg. <i>koža</i>	Russ. <i>koža</i>
*stj > *št'	*polh <sub>2</sub> - 'flat'	*plast'ь 'flat garment' later 'raincoat'	Pol. <i>plaszcz</i> , Cz. <i>plášť</i> , LSorb. <i>plášć</i>	Sln. <i>plašč</i> , BCMS <i>plašt</i> , Mac., Bulg. <i>plašt</i>	Russ., BRuss., Ukr. <i>plašč</i>
*zdj > *žd'	*dus-dyu- 'bad sky'	*дъзд'ь 'rain'	Pol. <i>deszcz</i> , Cz. <i>děšť</i> , Slk. <i>dážď</i> , ULSorb. <i>dešć</i>	Sln. <i>dež</i> , gen. sg. <i>dežja</i> , Sln. Prekmurje dial. <i>dešč</i> , <i>deždža</i> , Štok. arch. <i>dažd</i> , Mac. <i>dožd</i> , Bulg. <i>dažd</i>	Russ. <i>dožd'</i> , BRuss. <i>doždž</i> , Ukr. <i>došč</i>
*pj > *pl'	*plew- 'spit'	*pl'ujō, *pl'uva- 'spit' 1 sg., inf. stem	Pol., <i>pluje</i> , Cz. <i>pliji</i> , Slk. <i>pl'ujem</i>	Sln., BCMS <i>plujem</i> , Mac. <i>plue</i> , Bulg. <i>pljuja</i>	Russ. <i>plevat'</i> , BRuss. <i>pljavac'</i> , Ukr. <i>pljuvaty</i>
*bj > *bl'	*b <sup>h</sup> lewH- 'overflow'	*bl'ujō, *bl'uva- 'vomit' 1 sg., inf. stem	Pol. <i>bluję</i> , Cz. <i>bliji</i> , Slk. <i>bl'ujem</i> , ULSorb. <i>bluju</i>	Sln., BCMS <i>blujem</i> , Mac. <i>blue</i> , Bulg. <i>balva</i>	Russ., Ukr., <i>bljuju</i> , BRuss. <i>bljavac'</i>
*mj > *ml'	*d <sup>h</sup> ġemi 'earth' loc. sg.	*zeml'a 'earth'	Pol. <i>ziemia</i> , Cz. <i>země</i> , ULSorb. <i>zemja</i>	Sln., BCMS <i>zemlja</i> , Mac. <i>zemja</i> , Bulg. <i>zemja</i>	Russ., Ukr. <i>zemlja</i> , BRuss. <i>zjamlja</i>
*nj > *n'	(Gr. κυδωνια 'quince')	*dyn'a 'melon, pumpkin'	Pol. <i>dynia</i> , Cz. <i>dýně</i> , ULSorb. <i>dynja</i>	Sln. <i>dinja</i> , BCMS <i>dinja</i> , Mac., Bulg. <i>dinja</i>	Russ., BRuss., Ukr. <i>dynja</i>
*lj > *l'	*welh <sub>1</sub> - 'choose'	*vol'a 'will'	Pol., ULSorb. <i>wola</i> , Cz. <i>vůle</i> , Slk <i>vôla</i>	Sln., BCMS, Mac., Bulg. <i>volja</i>	Russ., BRuss., Ukr. <i>volja</i>
*rj > *r'	*ġher-/*ġhoro- 'shine'	*zor'a 'dawn, dusk'	Pol. <i>zorza</i> , Cz. <i>zoře</i> , Slk. <i>zora</i> , USorb. <i>zerja</i> , LSorb. <i>zorja</i>	Sln. <i>zarja</i> , BCMS, Mac. <i>zora</i> , Bulg. <i>zarja</i>	Russ., Ukr. <i>zarja</i> , BRuss. <i>zara</i>

Deiotation had also affected velars, with the results being identical to the first palatalization, e.g., PSI. \*<sub>DAWX-Y-Ā</sub> > OCS *duša* ‘soul’. The lack of the reflexes of the *j > l*’ change after labials in West Slavic and eastern South Slavic is a result of generalization of the non-iotated forms at the morpheme boundary. Toponyms reveal that the change had taken place throughout Slavic, e.g., Pol. *Grobla, Lublin*, Cz. *Liblín*, Bulg. *Koprivlen, Popovi Drǎvlja*.

PSI. \*w became *v* throughout Slavic, though there is a tendency to preserve the variant [w] in syllable-final position in some areas, e.g., BRuss. *low*, Sln. *lov* [low] < CSI. \**lovъ* ‘hunt, catch’.

TABLE 11.20 COMMON SLAVIC CONSONANT SYSTEM

	labial		dental		palatal		velar	
stops	<i>p</i>	<i>b</i>	<i>t</i>	<i>d</i>	<i>tʲ</i>	<i>dʲ</i>	<i>k</i>	<i>g</i>
fricatives	<i>f</i>		<i>s</i>	<i>z</i>	<i>š</i>	<i>ž</i>	<i>x</i>	( <i>ɣ</i> )
affricates			<i>c</i>					
nasals	<i>m</i>		<i>n</i>		<i>nʲ</i>			
tap/trill			<i>r</i>		<i>rʲ</i>			
liquid			<i>l</i>		<i>lʲ</i>			
glides	<i>v</i>				<i>j</i>			

## MORPHOLOGY

Slavic inflectional morphology continues the synthetic inflection type inherited from PIE, conserving case systems in the nominal system and the general patterns of inflection in the verbal system, albeit transformed by the syllable restructuring that affected the shape of endings. Of particular note is the development of the verbal aspect system in Slavic, whereby systematic distinctions of PERFECTIVE (“completed,” “viewed as a whole”) vs. IMPERFECTIVE hold for most verbs, e.g., OCS *kupiti* ‘buy’ perfect. : *kupovati* ‘buy’ impf. Sub-aspectual and lexical distinctions are accomplished by means of prefixation and suffixation, e.g., (PIE \**k<sup>(w)</sup>eg<sup>(h)</sup>-* >) OCS *čeznŏti, čeznŏ* ‘disappear’ inf., 1 sg., imperf. (CSI. \**jъz-čeznŏti*) > *išteznŏti* ‘disappear’ perfect.; *ištezati* ‘idem’ imperf. Within the verbs of motion in unprefixated verbs additional aspectual distinctions hold, often subsumed under the labels DETERMINATE (“unidirectional”) vs. INDETERMINATE (“manner of motion,” “multidirectional,” “iterative”), e.g., OCS *iti, nesti* ‘go’, ‘carry’ det., imperf.; *xoditi, nositi* ‘walk’, ‘carry’ indet., imperf.

Case marking is well preserved in all but eastern South Slavic – Mac., Bulg., and southern Serbian dialects, which were affected by contact phenomena in the Balkan convergence area. The inherited past tense systems were reduced in West and East Slavic as well as in Sln. and, to a more limited extent, BCMS, to a single construction based on the perfect (aux. of ‘be’ + *l*-participle). Mac. and Bulg. preserve only simple and coordinated tenses, but they also expanded to include evidentiality, “witnessed” vs. “non-witnessed,” e.g., Mac. *Toj beše vo Skopje* ‘he was in Skopje’ vs. *Toj bil vo Skopje* ‘he was supposedly/evidently in Skopje’ (Comrie & Corbett 1993: 272; though the synchronic category has been questioned, see Kempgen et al. 2009: 262–268).

## Nominal morphology

### Nouns

As in other PIE dialects nominal morphology has become restructured. In Slavic the productive declension patterns have fused PIE theme vowels with case and number markers

to make up the characteristic endings, while those in turn have been reshaped by phonetic changes in word-final position (see above regarding sound change). Root nouns of the type Lat. *nox, noctis* ‘night’ have generalized the non-nom. sg. case forms and reinterpreted the paradigm as an *i*-stem, e.g., OCS *noštъ* ‘night’ (← PSI. \**NAKT-I-M* acc. sg.). As a consequence of the rising sonority developments and the fusion of the vowels with grammatical markers, Slavic nominal morphology clearly distinguishes between lexical stems, which always end in a consonant, and grammatical endings, which consist of or begin with a vowel.

Consonantal stem types have largely merged with vocalic stem types in the daughter languages, although, curiously, Sln. has innovated by creating deverbative nouns from the relic *ū*-stem class, e.g., *osvežiti* ‘to refresh’ → *osvežitev* ‘refreshment’, perhaps as an extension of the tendency already observable in OCS to create deverbal nouns of the type *prěluby, prělubъve* ‘adultery’ nom., gen. sg. ← *prělubiti* (< *prě-* ‘trans’ + *lubiti* ‘love’), as well as conflation with deverbatives of the *-tъva* type, cf. Sln., BCMS, Bulg. *bitva* ‘battle’ ← CSI. \**biti* ‘to beat’. Of note is also the Slavic innovation of an *ent*-stem class denoting offspring, e.g., PIE \**d<sup>h</sup>eh<sub>1</sub>-* ‘suckle’ → OCS *děte, dětete* ‘child’ nom., gen. sg., PIE \**pork-* → OCS *prasę, prasęte* ‘piglet’ nom., gen. sg.

The nominal system of Slavic conserves all PIE cases but the ablative – which is thought to have merged with the genitive in the *o*-stem – as well as three genders (f., m., n.) and three numbers (sg., pl., du.). OCS and most of the daughter languages, with the exception of Mac. and Bulg., preserve six syntactic cases (nom., acc., gen., dat., loc., instr.) and a voc. form. OCS, Sln., and LSorb. retain dual number. With regard to form classes OCS displays the consonantal and vocalic stem classes of PIE, though some reorganization of the classes is already evident. The daughter languages have continued to reduce the stem classes, generally with the *o*- (*jo*-), *ā*- (*jā*), and *i*-stems emerging as the target paradigms, subsuming the unproductive consonantal stems. Grammatical gender predominantly aligns as follows: *o*- (*jo*-) = m., n.; *ā*- (*jā*) = f.; *i*- = fem; conflation of case forms occurs (most > least): du. > pl. > sg.; instr., loc., dat. > gen. > nom., acc.

### Vocalic stems

The productive *o*-stem masculines include older layers of inherited nouns, e.g., PIE \**sth<sub>2</sub>-lo-* > OCS *stolъ* ‘chair’, PIE \**mēmso-* > OCS *męso* ‘meat’, as well as *o*-grade deverbatives of the type OCS *vozъ* ‘wagon’ (cf. *vezъ* ‘I convey’), *plotъ* ‘fence’ (cf. *pletъ* ‘I weave’), and numerous derivatives with suffixes with the *o*-stem shape, e.g., OCS *ostrogъ* ‘palisade’ ← CSI. \**ostr-* ‘sharp’. The *u*-stem masculines generally merged in individual Late Common Slavic dialects with the *o*-stems, though in some of the daughter languages the endings have become part of the masculine declension type, sometimes with innovative functions, e.g., marking animacy in Cz. *Zbyňk-ovi* ‘to Zbyňek’ dat. sg. (vs. *stolu* ‘chair’ dat. sg.) and partitivity in Russ. *čaju* ‘some tea’ (vs. *cvet čaja* ‘the color of the tea’), the additional case meaning also being an effect of the Finnic substratum in northern Russian.

Innovative *-jo*- and *-jā*-stem declensions developed in Slavic, built on the *o*- and *ā*-stems, respectively. These differ mostly in the adjustment of intrasyllabic harmony, e.g., *jo*-stem \**MAR-I-AM* > \**mof-e* ‘sea’ nom., acc. sg. Inherited *jo*-stems are few, but the class remained productive and includes deverbatives such as OCS *vozъdb* ‘leader’ (cf. *voditi* ‘to lead’), as well as newer borrowings, e.g., OCS *kral’b* (from Germ. *Karl* ‘Charlemagne’). Innovative palatal stems, such as those that arose through the second palatalization of velars, also decline as *jo*-stems, e.g., OCS *pěņęđъb* ‘coin’ (from Germ. *penning*). The *jā*-stem type subsumed the relics of the \**-ih<sub>2</sub>* feminines, still attested by a

small number of nouns in OCS, e.g., PIE \*aldih<sub>2</sub> > OCS *ladii* ‘boat’, PIE \*peh<sub>2</sub>w- → OCS *pustyn’ii* ‘desert’ as well as in the f. nom. sg. form of the pres. act. ptcp., e.g., OCS *nesqšti* ‘carrying’ (< PSI. \*NESANT-Y-Ī).

TABLE 11.21 VOCALIC STEMS

class	PIE	OCS
<i>o</i> -stem	*g <sup>h</sup> ord <sup>h</sup> o- *mōyth <sub>2</sub> to-	<i>gradъ</i> ‘fortified place, city’ m. <i>město</i> ‘place’ n.
<i>jō</i> -stem	*mangyo- *polh <sub>2</sub> yo-	<i>mōžbъ</i> ‘man, husband’ m. <i>pol’e</i> ‘field’ n.
<i>u</i> -stem	*med <sup>h</sup> u-	<i>medъ</i> ‘honey’
<i>i</i> -stem	*nih <sub>2</sub> ti-	<i>nitъ</i> ‘thread’
<i>ā</i> -stem	*g <sup>w</sup> eneh <sub>2</sub>	<i>žena</i> ‘woman, wife’
<i>jā</i> -stem	*d <sup>h</sup> ghemy-eh <sub>2</sub>	<i>zemlja</i> ‘earth’

TABLE 11.22 CONSONANTAL STEMS

class	PIE	OCS
<i>n</i> -stem	*seh <sub>1</sub> men-	<i>sěmę</i> , gen. <i>sěmene</i> ‘seed’
<i>r</i> -stem	*d <sup>h</sup> ugh <sub>2</sub> ter-	<i>dъštĭ</i> , <i>dъštĕre</i> ‘daughter’
<i>s</i> -stem	*klewos-	<i>slovo</i> , <i>slovese</i> ‘word’
<i>nt</i> -stem	*ag <sup>w</sup> no-	<i>agnę</i> , <i>agnete</i> ‘lamb’
<i>ū</i> (UW)-stem	*swekruh <sub>2</sub> -	<i>svekry</i> , <i>svekrъve</i> ‘husband’s mother’

### Nominal declension

In the following chart the presumed PIE input is compared with the attested OCS endings (CSl., where *ъ* occurs before *j*, which yields *ij* in OCS), which are presumed to be similar to those of LCSl. Wherever the arrow → is used, a non-linear development is presumed to have occurred. The trajectories of the endings, as well as the PIE input, are a subject of much debate, and the detail needed to explain them far exceeds the scope of this chapter. An emblematic issue is the origin of the *-ъ* ending in the nom. and acc. sg. cases of the *o*-stems. The expected outcome, following the regular phonetic development, loss of final consonants, would have been nom. sg. \*-os > *-o*. (This outcome is in fact represented regionally in proper nouns, where the ending may have earlier been interpreted as vocative or hypocoristic, BCMS *Branko*, *Vlado* [← *Branislav*, *Vladimir*]; Ukr. *Danylo*, *Petro*. In the Finnic substratum areas of Pskov-Novgorod, as attested in the Birchbark letters, the vocative ending *-e* was generalized, on which see Vermeer 1994.) Such an outcome threatened merger with the neuter *o*-stems, as well as neutralizing the distinction between agent and patient roles. Among alternatives – other solutions have been considered – was to select the *u*-stem endings in *-ъ*, which did not solve the problem of agent-patient distinctions. This problem was repaired by the development of the category of animacy, whereby genitive case marking was introduced for animate referents in the accusative position, e.g., OCS *vъsi bo ěko proroka imōtъ Ioana* (Matthew 21:26, Codex Marianus) ‘all thus as a prophet-ANIM. take John-ANIM.’ The input of the instr. sg. and dat., instr. pl. and du. forms assume a PIE dialect input of \*-m- (as opposed to \*-b<sup>h</sup>-), common to Germanic, Baltic, and Slavic. The schema below shows the expected PIE input on the left and CSl. on the right (with PSI. for illustration in some instances), where > indicates a direct phonetic continuation and → indicates a post-PIE rearrangement of the material (PIE input, with some adjustments, based on Sihler 1995: 248).

TABLE 11.23 VOCALIC STEM ENDINGS

	<i>o</i> -stems	<i>u</i> -stems	<i>i</i> -stems	$\bar{a}$ -stems
<i>singular</i>				
nom.	*-os → * $\bar{u}$ > - $\bar{b}$	*-us > - $\bar{b}$	*-is > - $\bar{b}$	*-eh <sub>2</sub> > - <i>a</i>
(n.)	*-om > <i>o</i>			
acc.	*-om → * $\bar{u}$ > - $\bar{b}$	*-um > - $\bar{b}$	*-im > - <i>b</i>	*-eh <sub>2</sub> m > - $\varphi$
gen.	*-oh <sub>2</sub> (e)t (abl.) > - <i>a</i>	*-ows > - <i>u</i>	*-eys > - <i>i</i>	*-eh <sub>2</sub> os → * $\bar{u}$ > - <i>y</i>
dat.	*-ōy → * $\bar{a}$ w > - <i>u</i>	*-ewey > - <i>ovi</i>	*-eyey > * $\bar{e}$ y > - <i>i</i>	*-eh <sub>2</sub> ey > * $\bar{a}$ y > - $\bar{e}$
instr.	*-oh <sub>1</sub> → * $\bar{a}$ MI > -om $\bar{b}$	*-umi > - $\bar{b}$ m $\bar{b}$	*-imi > - $\bar{b}$ m $\bar{b}$	*-eh <sub>2</sub> mi → * $\bar{a}$ YAM > - <i>o</i> j $\varphi$
loc.	*-oy > - $\bar{e}$	*-ēw > - <i>u</i>	*-ēy > - <i>i</i>	*-eh <sub>2</sub> y > * $\bar{a}$ y > - $\bar{e}$
voc.	*-e > - <i>e</i>	*-ew > - <i>u</i>	*-ey > - <i>i</i>	*-h <sub>2</sub> > - <i>o</i>
<i>plural</i>				
nom.	*-oy > - <i>i</i>	*-ewes > - <i>ove</i>	*-eyes > - <i>bje</i>	= acc. pl.? > <i>y</i>
nom., acc. n.	*-eh <sub>2</sub> > - <i>a</i>			
acc.	*-oms > - <i>y</i>	*-ums > - <i>y</i>	*-ims > - <i>i</i>	*-eh <sub>2</sub> ms > - <i>y</i>
gen.	*-ōm → * $\bar{u}$ > - $\bar{b}$	*-owom > - <i>ov<math>\bar{b}</math></i>	*-eyom > - <i>bj<math>\bar{b}</math></i>	*-eh <sub>2</sub> ōm > - $\bar{b}$
dat.	*-omus → - <i>om<math>\bar{b}</math></i>	*-umus → - $\bar{b}$ m $\bar{b}$	*-imus > - $\bar{b}$ m $\bar{b}$	*-eh <sub>2</sub> mus > - <i>am<math>\bar{b}</math></i>
instr.	*-ōys → * $\bar{u}$ > - <i>y</i>	*-umīs > - $\bar{b}$ mi	*-imī > - $\bar{b}$ mi	*-eh <sub>2</sub> miHs > - <i>ami</i>
loc.	*-oysu > - $\bar{e}$ x $\bar{b}$	*-usu > - $\bar{b}$ x $\bar{b}$	*-isu > - $\bar{b}$ x $\bar{b}$	*-eh <sub>2</sub> su > - <i>ax<math>\bar{b}</math></i>
<i>dual</i>				
nom., acc.	*-eh <sub>2</sub> > - <i>a</i>	*-uh <sub>2</sub> > - <i>y</i>	*-ih <sub>2</sub> > - <i>i</i>	*-eh <sub>2</sub> y > - $\bar{e}$
nom., acc. n.	*-h <sub>2</sub> y > - $\bar{e}$			
gen., loc.	*-h <sub>2</sub> ow > - <i>u</i>	*-owh <sub>2</sub> ow > - <i>ovu</i>	*-eyh <sub>2</sub> ow > - <i>bj<math>\bar{u}</math></i>	*-eh <sub>2</sub> ow > - <i>u</i>
dat., instr.	*-omoH > - <i>oma</i>	*-umoH > - $\bar{b}$ ma	*-imoH > - $\bar{b}$ ma	*-eh <sub>2</sub> moH > - <i>ama</i>

TABLE 11.24 CONSONANTAL STEM ENDINGS	
<i>singular</i>	
nom.	*-s (m.), *- $\emptyset$ > - $\emptyset$ (implies Slavic stem changes)
acc.	*- $\bar{m}$ > - $\emptyset/\bar{b}$
gen.	*-es > - <i>e</i>
dat.	*-ey > - <i>i</i>
instr.	*? > * $\bar{a}$ /EMI → - $\bar{b}$ /em $\bar{b}$ (m.); *? > * $\bar{a}$ YAM > - <i>bj<math>\varphi</math></i> (f.)
loc.	*-i → - <i>i/e</i>
voc.	= nom. sg.
<i>plural</i>	
nom.	*-es > - <i>e</i> ; - <i>i</i> , - <i>a</i> from <i>o</i> -stems
acc.	*- $\bar{m}$ s > - $\bar{b}$ ; - <i>i</i> , - <i>a</i> , - <i>y</i> from <i>o</i> -, $\bar{a}$ -stems
gen.	*-om > - $\bar{b}$ from <i>o</i> -stems?
dat.	*-(i)mus > - $\bar{b}$ m $\bar{b}$
instr.	*-(i)mī > - $\bar{b}$ mi; - <i>y</i> from <i>o</i> -stems
loc.	*-(C)su → - $\bar{b}$ x $\bar{b}$ /-ix $\bar{b}$ /-ex $\bar{b}$
<i>dual</i>	
nom., acc.	*-ih <sub>1</sub> > - <i>i</i> (r-stems, <i>db<math>\bar{b}</math></i> ); - <i>a</i> , - $\bar{e}$ from <i>o</i> -, $\bar{a}$ -stems
gen., loc.	*-ow > - <i>u</i>
dat., instr.	*-imoH > - $\bar{b}$ ma

### Adjectives

Adjectival declension derives from nominal declensions, where agreement in gender, number, and case follows the pattern of the *o-*, *jo-* stems for masculine and neuter and the *ā-*, *jā-* stems for feminine, PIE \*new-os, \*new-om, \*new-eh<sub>2</sub> > OCS *nov-ъ*, *nov-o*, *nov-a* ‘new’ nom. sg. m., n., f.; Gmc \*tewd- → PSl. \*TEWD-Y- > CSL. \*t’ud’- > OCS *štužd-b*, *štužd-e*, *štužd-a* ‘foreign’ (with variant base forms *tužd-*, *stužd-*, perhaps as a function of dissimilation). Of particular note is the development of an articulated form of the adjective with \*j- pronouns (see below), which mark definiteness, an innovation shared with Baltic, e.g., *novъ-jb*, *novо-je*, *nova-ja*; *nova-jego* gen. sg. m./n., *novu-jemu* dat. sg. m./n., etc. Cf. *naujas* ‘new’ m. sg. indef. vs. *naujasis* def. (< *naujas* + *jis*).

Comparative forms are produced with *-jb/-bš-* or innovative *-ěj-bš-* (shortened in the nom. sg. m. and n. by word-final consonant elision) and follow the *j-* declension types, e.g., OCS *bol’i*, *bol’e*, *bol’bša* ‘bigger’; OCS *nověi*, *nověj(j)e*, *nověiša* ‘newer’.

Participial forms also belong to the adjectival declensions, e.g., present active participle: OCS *nesy* (m., n.), *nesošti* (< PSl. \*NESA-NT-S, \*NESA-NT-M, \*NESANT-Y-I) ‘carrying’; past active participle *nesъ* (m., n.), *nesъši* (< PSl. \*NES-US-S, NES-US-M, \*NES-US-Y-I) ‘having carried’; past passive participle *nesenъ*, *neseno*, *nesena* (< PSl. \*NES-EN-AS, -A, -Ā) ‘having been carried’; present active participle *nesomъ*, *nesomo*, *nesoma* (< PSl. \*NES-AM-AS, -A, -Ā) ‘being carried’. The resultative “*l*-participle” occurs only in the nom. forms, e.g., *neslъ*, *neslo*, *nesla*, which occur in compound tenses and mood constructions.

### Pronouns

Personal pronouns in Slavic conserve much of the PIE pattern, distinguishing six cases as in the noun. Here the OCS examples stand (nearly) for all of Slavic. PIE \*h<sub>1</sub>eǵ, \*h<sub>1</sub>eǵHom > Cz. *já*, Pol., Slk., ULSorb., BCMS, BRuss., Ukr., Russ. *ja* || OCS *azъ*, Sln. *jaz* ‘I’, Mac. *jas*, Bulg. *az* (with prothetic *j-* and lengthening by Winter’s Law; on the heterogeneous inheritances see Kapović 2009); PIE \*h<sub>1</sub>mene > OCS *mene*, *mę* ‘me’ – full and clitic forms; PIE \*tuH > OCS *ty* ‘you’ 2 sg. nom.; in OCS *tebe*, *te* 2 sg. acc., the expected stem from PIE \*tewe does not occur but is replaced by the stem from the dat. sg., \*tebh- (OCS *tebě* 2 sg. dat. sg.); the nom. (and acc.) forms of the 1 and 2 pl. are evidently reshaped on the model of *ty*: OCS *my* ‘we’ nom., *ny* ‘us’ acc., *vy* ‘you’ (pl. and honorific) nom., acc. The acc., gen. forms generalized the lengthened vowel root and must have been extended by a suffix, PIE \*nōs, \*wōs → PSl. \*NĀS-A(M), \*WĀS-A(M) > OCS *nasъ*, *vasъ*. The homophonic loc. pl. forms OCS *nasъ*, *vasъ* provide the direct continuation of the loc. pl. ending in *-su*, which has been generalized in the noun and adjective paradigms as *-ъ* (following the RUKI change). The reflexive pronoun follows the pattern of the 3 sg.: OCS *sę* acc., *sebe* gen., *sebě*, *si* dat. full and clitic, *sebě* loc., *soboję* instr. Third person pronouns, as well as corresponding relatives, are built from the anaphoric pronoun PIE \*yo- apart from the nom. forms, e.g., OCS *jb* acc. sg. m., *je* acc. sg. n., *ję* acc. sg. f., *jego* gen. sg. m., n.; *jeję* gen. sg. f.; *jemu* dat. sg. m., n., *jejb* dat. sg. f. In prepositional constructions, initial *j* is replaced by *n’* (< \*nj) as a result of the reanalysis of *-n* in prepositions, \*VUN ‘in’, \*SUN ‘with’ (< PIE \*h<sub>1</sub>ŋ, \*kon) as *n-*, e.g., *otъ n’ego* ‘from him, it’, *kъ n’ejb* ‘toward her’, *sъ n’imъ* ‘with him’. In OCS the nom. forms from anaphoric \*yo- occur only in relativizers compounded with the particle (focus marker) *že* (the o-grade of which appears in the gen. sg. m. and n. pronominal and adjectival ending: *jego*), e.g., *čto že vidiši sočъsbъ jъže estъ vъ očese bratra toego a brvno eže estъ vъ očese tvoemъ ne čuješi* ‘what you-see a mote-acc.-sg.-m. that-nom.-sg.-m. is in the eye of brother yours

but the log-acc.sg.-n. that-nom.sg.-n. is in eye yours not you-feel' (Luke 6:41, Codices Zographensis, Marianus). In the daughter languages the usual 3 pers. nom. pronouns are built from CSI. \*on-, e.g., OCS *онѣ, она, оно; oni, ony, ona* m., f., n. sg., pl. (< PIE \*h<sub>2</sub>en-), though Mac. and Bulg. use forms built from PIE \*to-, e.g., Bulg. *toj, tja, to* m., f., n. sg. nom. In other Slavic languages, pronouns built on this form usually serve a distal, switch-reference, or resumptive function.

TABLE 11.25 OCS 3 PERS. PRONOUN

	sg.	pl.	du.
nom.	*jъ m., *ja f., *je n.	*ji m., *jĕ f., *ja n.	*ja m., *ji f., n.
acc.	jъ m., jŏ f., je n.	jĕ m., f.; ja n.	= nom.
gen.	jego m., n.; jĕjĕ f.	ixъ	jeju
dat.	jemu m., n.; jĕjъ f.	imъ	ima
loc.	n'emъ m., n.; n'ejъ f.	ixъ	jeju
instr.	imъ m. n.; jĕjŏ f.	imi	ima

OCS and the Slavic daughter languages preserve other 3 pers. pronouns marked for proximity in time or space, e.g., OCS *сѣ, си, се* 'this here' (< PIE \*ki-); BCMS *ovaj, ova, ovo* 'this' (< PIE \*h<sub>2</sub>ew-) – contrasted with *onaj, ona, ono* 'that', cf. Sln. *one* 'what-chamacallit' < CSI. \*ono + \*je). The deictic *s-* particle also plays a role in the formation of adverbs of time, OCS *дѣньсѣ*, Cz., Slk. *dnes*, BCMS *danas* || Russ. *segodnja* 'today' (< CSI. \*дѣнь + сѣ || \*sego + дне); OCS *си ношти*, BCMS *sinoć*, Mac. *sinoka* 'last night' (< CSI. \*si not'i).

Interrogative pronouns in Slavic display both conservative and innovative features, as is evident from the following table. The nom. forms of both the animate and inanimate have been extended with the element \*to(-s, -d), though variant forms also occur, e.g., Pol. *nikt*, Sln. *nihče* 'nobody' < CSI. \*nikъtъ(-že); the littoral (Čakavian) dialect of Croatian has the bare form of the inanimate *ča* (< CSI. \*čъ). Sln. and Croatian Kajkavian dial. *kaj* 'what' is formed from a different inherited element, PIE \*k<sup>w</sup>eh<sub>2</sub>, extended with the anaphoric element CSI. \*jъ < PIE \*yo- (Snoj 2003: s.v.). The same element is found as a subordinating conjunction in Bulg., Pol. dial., as well as in Prekmurje Sln. *ka*, where it contrasts with irrealis *da* (< PIE \*doh<sub>2</sub>), the (otherwise indicative) prevailing subordinating conjunction in western South Slavic (see Greenberg 2011).

TABLE 11.26 INTERROGATIVE PRONOUNS

	PIE	OCS	Cz.	Sln.	Russ.					
nom.	*k <sup>w</sup> o- 'who'	*k <sup>w</sup> id 'what'	<i>кто</i>	<i>čto</i>	<i>kdo</i>	<i>co</i>	<i>kdo</i>	<i>kaj</i>	<i>kto</i>	<i>čto</i>
acc.	*k <sup>w</sup> om	*k <sup>w</sup> im	= gen. = nom.	= gen. = nom.	= gen. = nom.	= gen. = nom.	= gen. = nom.	= gen. = nom.	= gen. = nom.	= nom.
gen.	*k <sup>w</sup> oso	*k <sup>w</sup> eso	<i>kogo</i>	<i>česo</i>	<i>koho</i>	<i>čeho</i>	<i>koga</i>	<i>česa</i>	<i>kogo</i>	<i>čego</i>
dat.	*k <sup>w</sup> osmōy	*k <sup>w</sup> esmey	<i>komu</i>	<i>česomu</i>	<i>komu</i>	<i>čemu</i>	<i>komu</i>	<i>čemu</i>	<i>komu</i>	<i>čemu</i>
loc.	*k <sup>w</sup> osmi	*k <sup>w</sup> esmi	<i>комъ</i>	<i>čеть</i>	<i>kom</i>	<i>čem</i>	<i>kom</i>	<i>čem</i>	<i>kom</i>	<i>čem</i>
instr.	*k <sup>w</sup> Vh <sub>1</sub> -	*k <sup>w</sup> Vh <sub>1</sub> -	<i>кѹтъ</i>	<i>čить</i>	<i>kým</i>	<i>čím</i>	<i>kom</i>	<i>čim</i>	<i>kem</i>	<i>čem</i>

Further pronominal bases include PIE \*ēyno- > OCS *иъ* 'other', PIE \*ēysth<sub>2</sub>o- > OCS *истъ* 'same', PIE \*wik- > OCS *всѣ* 'all'.

## Indeclinables

Adverbs of time, place, manner, and direction were created by combining pronominal elements with deictic particles.

TABLE 11.27 OCS ADVERBS OF PLACE, TIME, MANNER, DIRECTION

	PSl. *-d <sup>h</sup> e ‘where’	PSl. *-DĀ, -KUDĀ ‘when’	PSl. *-(Ā)KA ‘how’	PSl. *-ĀMA ‘direction’
*k-	<i>kъde</i> ‘where’	<i>kъda, kogda</i> ‘when’	<i>kako</i> ‘how’	<i>kamo</i> ‘whither’
*s-	<i>sъde</i> ‘here’	<i>sъda</i> ‘now’	<i>sice</i> (< *SĪKA) ‘in this way’	<i>sĕmo</i> ‘hither’
*t-		<i>togda</i> ‘then’	<i>tako</i> ‘thus’	<i>tamo</i> ‘there’
*ov-	<i>ovъde</i> ‘there’	<i>ovogda</i> ‘then’		<i>ovamo</i> ‘thither’
*on-	<i>onъde</i> ‘over there’	<i>onъda</i> ‘at that time’		<i>onamo</i> ‘yonder’
*j-	<i>jъde(že)</i> ‘at that place’	<i>jeda</i> ‘at the time when’	<i>jako</i> ‘thus’	
*vъs-	<i>vъsъde</i> ‘everywhere’	<i>vъsegda</i> ‘always’	<i>vъsako</i> ‘in all ways’	

Other adverbs include OCS *tu* ‘here’ (< PIE \*tow), *mъnogo* ‘much’ (< PIE \*mnog<sup>h</sup>o-). Productive adverbial formants include *-o*, *-ě*, *-bъsky*, e.g., OCS *dobro*, *dobrĕ*, both ‘well’; *slovĕnъsky* ‘Slavic’.

## Verbal morphology

The Slavic verb materially reflects a simplified picture of Indo-European verbal morphology, though, as mentioned above, Slavic has innovated in its morphologized aspectual system by means of prefixation and suffixation as well as by extending the ablaut pattern. Only one example of PIE reduplication persists in the verb, and this is in the small group of athematic verbs, i.e., OCS *dadĕtъ* ‘they give’ (PSl. \*DĀDEN-T- ← PIE \*deh<sub>3</sub>-, cf. Skr. *dadāmi* ‘I give’). The remaining athematic verbs in Slavic are (PIE \*h<sub>1</sub>es- >) OCS *jesmъ*, *jesi*, *sъtъ* 1 sg., 2 sg., 3 pl. (inf. *byti*) ‘be’; (PIE \*weyd- >) *vĕmъ* 1 sg. (inf. *vĕdĕti*) ‘know’ (an isolated relic form descended from the PIE intransitive conjugation is found in OCS, OCz., ORuss. *vĕdĕ*, on which see Ivanov 1981: 68); (PIE \*ed- → PSl. ĒD-MI >) OCS *ĕmъ/jamъ* (inf. *ĕsti, jasti*). The verb *imĕti*, 1 sg. *imamъ* ‘to have’ (< CSI. \*jъmamъ < PSl. \*IM-Ā-MI ← PIE \*ṛ- ‘grasp, take’), also belongs here – it is related to an *e/o*-thematic verb *ĕti, jъmъ* ‘to take’ inf., 1 sg.

The present tense endings, as in the nominal declension, have been reorganized. The 1 sg. ending of the thematic type in *-o* cannot have been original but was built from the inherited PIE \*-oh<sub>2</sub>, and the *-m* marker was extended from the athematics after the loss of final consonants. The extension of the athematic 1 sg. *-m* made a second comeback in the historical period with the rebuilding of the present tense markers in West Slavic and western South Slavic, e.g., Slk. *nesiem*, Sln. *nesem*; Pol. *gram*, Cz. *hrám*, Sln., BCMS *igram* || Russ. *igraju*, Ukr. *hrajju*, Bulg. *igraja*, Plb. *jaigroja* (< CSI. \*jъgrajō) ‘play’ 1 sg. (see further Janda 1996). The 3 sg. and pl. forms in OCS *-tъ* are of secondary origin, most likely from the 3 sg. pronoun in *t-*. East Slavic, however, preserves the original athematic ending in *-ti*, e.g., ORuss. *jestъ* ‘is’, Russ. *est’*, BRuss. *ĕsc’*, albeit with considerable dialectal variation (see Miller 1988 for details). A zero 3 sg. ending (inherited from PIE) is also amply attested, e.g., Cz. *nese*, Slk. *nesie*, Sln. *nese*, Ukr. *nese*. Similarly, 1 pl. endings show variation through generalization of the ablaut variants *-me/mo*, most likely in response to competition from the 1 sg. marker in *-m*, cf. USorb. *sym – smy*, Cz. *jsem* ‘I



amʹ – *jsme* ‘we are’, Slk. *som* – *sme*, Sln. *sem* – *smo*, BCMS *sam* – *smo*, Mac. *sum* – *sme*, Bulg. *səm* – *sme*.

TABLE 11.28 INFLECTION OF NON-PAST FINITE VERB FORMS

	athematic	thematic
1 sg.	<i>damb</i> (< *PSI. *DĀD-MI)	<i>nesq</i> (< *NES-Ā-M)
2 sg.	<i>dasi</i> (< *DĀD-SI)	<i>neseši</i> (< *NES-E-XĪ)
3 sg.	<i>dastb</i> (< *DĀD-T-)	<i>neseṭb</i> (< *NES-E-)
1 pl.	<i>damb</i> (< *DĀD M-)	<i>nesemṭb</i> (< *NES-E-M-)
2 pl.	<i>daste</i> (< *DĀD-TE)	<i>neseṭe</i> (< *NES-E-TE)
3 pl.	<i>dadęṭb</i> (< *DĀD-EN-T-)	<i>nesqṭb</i> (< *NES-A-NT-)
1 du.	<i>dasta</i> (< *DĀD-TĀ)	<i>neseṭa</i> (< *NES-E-T-Ā)
2 du.	<i>daste</i> (< *DĀD-TE)	<i>neseṭe</i> (< *NES-E-TE)
3 du.		

Traditionally, OCS verbs are grouped into five classes, following the classification by Leskien 1969 (1871), which corresponds to layers of archaism vs. innovation, and are indexed by their characteristic theme (*e/o*, *ne/o*, etc.). Two basic stems are accounted for, represented by the infinitive, which normally contains a (semantically empty) suffix, e.g., LCSl. \*bъr-a-ti ‘to take’, and a present or non-past-tense stem \*ber-e- ‘takes’. Sub-types are determined by the suffix or lack of it (in which case the stem-final segment is relevant) in the infinitive stem. In the table below the OCS forms are given in the inf., 1 sg., 3 sg. pres. tense (the listing of types here is not comprehensive).

TABLE 11.29 VERB CLASSES, FOLLOWING LESKIEN

class	theme	PIE	PSl.	OCS
I.A	<i>e/o</i>	*h <sub>1</sub> nek-	*NES-TĒY *NES-Ā-M *NES-E-	<i>nesti</i> ‘carry’ det. <i>nesq</i> <i>nesetb</i>
1.B		*b <sup>h</sup> er-	*BIR-Ā-TĒY *BER-Ā-M *BER-E-	<i>bъrati</i> ‘take’ <i>berq</i> <i>beretb</i>
II	<i>ne/o</i>	*steh <sub>2</sub> -	*STĀ-TĒY *STĀ-NĀ-M *STĀ-NE-	<i>stati</i> ‘stand up’, ‘begin to’ <i>stanq</i> <i>stanetb</i>
		*kos-	*KAS-NŪ(N)-T-EY *KAS-NĀ-M *KAS-NE-	<i>kosnqi</i> ‘touch’ <i>kosnq</i> <i>kosnetb</i>
III.1.A	<i>je/o</i>	*ǵnoh <sub>3</sub> -	*ZNĀ-TĒY *ZNĀ-YĀ-M *ZNĀ-YE-	<i>znati</i> ‘know’ <i>znajq</i> <i>znajetb</i>
III.1.B		*kes-	*KES-Ā-TĒY *KES-YĀ-M *KES-YE-	<i>česati</i> ‘gather (fruit)’ <i>češq</i> <i>češetb</i>
III.2.B		*(s)newH-	*AB-SNAW-Ā-TĒY *AB-SNAW-YĀ-M *AB-SNAW-YE-	<i>osnovati</i> ‘found’ <i>osnujq</i> <i>osnujetb</i>
IV.A	<i>ī</i> (< PIE *-eye-)	*h <sub>1</sub> nok-	*NAS-EY-TĒY *NAS-Y-Ā-M *NAS- I(Y)E-	<i>nositi</i> ‘carry’ indet. <i>nošq</i> <i>nositb</i>

class	theme	PIE	PSl.	OCS
IV.B		*sed-	*SĒD-Ē-TĒY *SĒD-Y-Ā-M *SĒD-I(Y)E-	<i>sěděti</i> ‘sit’ stative <i>sěždō</i> <i>sěditb</i>
IV.B		*b <sup>h</sup> oyH-	*BAY-Ē-TĒY *BAY-Ā-M *BAY-EY-	<i>bojati sę</i> ‘be afraid’ <i>bojō sę</i> <i>bojiti sę</i>
V	athematic	*b <sup>h</sup> uh <sub>2</sub> - *h <sub>1</sub> es-	*BŪ-TĒY *ES-MI *ES-TI	<i>byti</i> ‘to be’ <i>esmb</i> <i>estb</i>
V		*woyd-	*WAYD-Ē-TĒY *WAYD-MI *WAYD-TI	<i>věděti</i> ‘know, be aware’ <i>věmb</i> <i>věstb</i>

Form classes and compound tenses/moods of the verb, as attested in OCS, include the following:

TABLE 11.30 SLAVIC VERB FORM CLASSES

infinitive	<i>nesti</i> ‘to carry’
supine	<i>nestb</i> ‘(to go) carry’
present tense	<i>nesō, neseši</i> ‘carry’ 1 sg., 2 sg.
aoist	<i>něxъ, nese</i> ‘carried’ 1 sg., 2, 3 sg.
imperfect	<i>nesěaxъ, nesěaše</i> ‘was carrying’ 1 sg., 2, 3 sg.
imperative, optative	<i>nesi, nesěte</i> ‘carry’ 2, 3 sg., 2 pl.
present active participle	<i>nesy, nesōšti</i> ‘carrying’ nom. sg. m./n., nom. sg. f.
past active participle	<i>nesъ, nesъši</i> ‘having carried’ nom. sg. m./n., nom. sg. f.
present passive participle	<i>nesomb, nesomo, nesoma</i> ‘being carried’ m., n., f. sg.
past passive participle	<i>nesemb, neseno, nesena</i> ‘having been carried’ m., n., f. sg.
perfect	<i>estb nesb</i> ‘has carried’ m. sg.
pluperfect	<i>bě nesb, běaše nesb</i> ‘had carried, had been carrying’
conditional	<i>bi nesb</i> ‘would carry’ 2, 3 sg.
future anterior	<i>bōdetb nesb</i> ‘will have carried’ 3 sg.

### *Ablaut, tense, and aspect*

Slavic tenses and their corresponding form classes have become simplified from what was a much more complex earlier system in late PIE, with partially differing results in Baltic and Slavic. Slavic and Baltic inherited *e*-grade thematic verbs (PIE \*wed<sup>h</sup>- > OCS *vedetb* ‘leads’ det., Lith. *veda*; \*g<sup>w</sup>eyh<sub>3</sub>- OCS *žiti, živō, Cz. žít, žiji* ‘live’ inf., 1 sg.) as well as derived *o*-grade causatives, iteratives, and deverbal nouns (OCS *voditi* ‘to lead’ indet., Lith. *vadýti* ‘idem’, OCS *voždъ* ‘leader’, *vojevoda* ‘warlord’; PIE \*g<sup>w</sup>oyh<sub>3</sub>- > ORuss. *goiti* ‘revive’, Cz. *hojit* ‘heal’, Sln., BCMS *gojiti* ‘raise [a child]’). On the other hand, Slavic possesses sets of synthetic past tense forms (“aoist” and “imperfect”) built from PIE \*-s, which is lacking in Baltic, whereas Baltic possesses an *s*-future lacking in Slavic (with the exception of the relic pres. act. ptcp. form OCS *byšęšt-/byšōšt-* [< PSl. \*BŪ-S-Y-E/ANT-Y-] ‘that which is to be’, replaced by *bōdōšt-* ‘idem’, contrasting with *sōšt-* ‘being’; see Vaillant 1966 [=1950–77]); see also Darden 1990, Andersen 2013. Lacking a separate synthetic future form, the primary Slavic future is expressed by the same form as for the present tense, which could be termed “prospective/actual” (following Andersen 2013),

where in the daughter languages perfective verbs in the non-past tend to be understood as future; in ESL perfective verbs in the non-past forms are exclusively understood as future. Compound future tenses, in addition to the future anterior attested in OCS *boǰetъ nesla* ‘will have carried’, which has become the generalized future construction in Sln. and Kajkavian Croat. (Sln. *bo nesla* ‘will carry’) and one of the future constructions in Pol. *będę niosła* or *będę nieść* (both with imperf. verbs only), developed in the daughter languages, e.g., BCMS *pisat će* ‘will write’, Bulg. *šte piša* ‘will write’ (both auxiliaries from CSI. \*(xo)tǃęti ‘want’), Ukr. *budu pysati* ‘I shall write’ or *pysatymu* (< CSI. \*pisati + \*jǃmǃ) ‘idem’ (both with imperf. only).

Slavic extended the pattern of ablaut as a means of (partially) articulating innovative aspectual morphology, in effect reinvigorating the vřddhi grade, to produce imperfective partners for new perfective lexical items produced through prefixation, e.g., (PIE \*leg<sup>h</sup>- ‘lie [down]’ >) OCS *lešti, lęǰǃ* ‘lie down’ inf., 1 sg. perfect., *položitǃ* ‘place something standing’ perfect., *polagati, polagajǃ* ‘idem’ imperf. (PSI. \*-LEG-/LAG- → -LĀG-). The innovation builds also on the excrescent vowels that developed from syllabic sonorants, e.g., OCS *sǃbǃrati* ‘collect’ perfect., *sǃbirati* ‘idem’ imperf. (PIE \*b<sup>h</sup>ǃ- > PSI. \*BIR- → \*BĪR-). The rise of the aspectual system of Slavic as such is still being debated, with significant differences of opinion. Recent work indicates that the system of verbal aspect developed through heterogeneous developments that coalesced into a general schema opposing perfective and imperfective aspects. Among the processes were, in addition to the ablauting pattern just mentioned, (1) the merger of synthetic imperfect and aorist into a single preterite in PSI., which resulted in the attested OCS aorist, and the rise of a new agglutinative synthetic preterite, the imperfect (see above) (see Andersen 2013); (2) prefixation of simplex verbs, creating new lexical meanings, *aktionsarten*, as well as purely aspectual distinctions, “*préverbe vide*” in the perfective aspect (e.g., OCS *ubiti* ‘kill’ inf. perfect. < PIE \*ow- ‘ablative’ + \*b<sup>h</sup>eyH- ‘beat’; OCS *uspęti* ‘succeed’ inf. perfect. < PIE \*ow- + \*sp<sup>h</sup>eh<sub>1</sub>- ‘flourish’) (see Dickey 2005 and forthcoming); and (3) suffixation, allowing the articulation of imperfective partners for lexicalized prefixed perfectives (OCS *ubijati, ubivati* ‘kill’ inf. imperf.; *uspęvati* ‘succeed’ inf. imperf.). The general meanings of the perfect. : imperf. opposition in the daughter languages differ geographically, with two clear-cut sub-types: “eastern” (ESL. + Bulg.) and “western” (Cz., Slk., ULSorb. + Slovene), as well as two transitional types, Pol. (trending toward “eastern”) and BCMS (trending toward “western”). In each type the perfect. is marked for a general meaning, whereas the other imperf. is unmarked as such: the “eastern” type conveys “temporal definiteness” (e.g., succession of events) and the “western” type “totality” (for details, see Dickey 2000).

## WORD FORMATION

Slavic lexis has considerably innovated new configurations of PIE material, including ‘hand’ (Lith. *ranka*) < PIE \*wronk-eh<sub>2</sub> ← \*wrenk- ‘grasp’; OCS *noga* ‘foot, leg’ (Lith. *naga* ‘hoof’) ← PIE \*h<sub>3</sub>nogh<sup>o</sup>- ‘(toe)nail’ + collective suffix \*-eh<sub>2</sub>, cf. OCS *nogǃbъ* ‘(finger/toe)nail’ (Lith. *nagutis* ‘idem’). Slavic word formation makes considerable use of prefixation and suffixation, as well as, in regard to the verb, extensions of ablaut patterns, as was mentioned above in the discussion of the verb.

**TABLE 11.31 DEVERBATIVES FROM CSL. \*VĚDĚTI ‘TO KNOW’, DERIVATIVES OF CSL. \*RYBA ‘FISH’**

CSL.	Attestation	Meaning
*věd-a	Cz. <i>věda</i>	‘science, knowledge’
*ne-věd’-a	OCS <i>nevěžda</i>	‘idiot’
*ne-věd’-ь	OCS <i>nevěždb</i>	‘idiot’
*ne-věd’-ьstvo	OCS <i>nevěžьstvo</i>	‘ignorance’
*ne-věd-om-ь-jь	Cz. <i>nevědomý</i>	‘unknown’
*ne-věd-ьnь	Sln. <i>neveden</i> , BCMS arch. <i>nev(j)edan</i>	‘ignorant’, ‘arrogant’
*ryb-ar’-ь	OCS <i>rybar’b</i>	‘fisherman’
*ryb-ar’-i-ti	Cz. <i>rybařit</i> , Sln. <i>ribariti</i>	‘to fish’
*ryb-ic-a (< PSl. *RŪB-ĪK-Ā)	OCS <i>rybica</i>	‘small fish’
*ryb-ьn-ь-jь	Russ. <i>rybnoj</i>	‘pertaining to fish’
*ryb-ьn-ik-ь	Cz. <i>rybník</i>	‘pond’
*ryb-ьn-ic-a	Pol. <i>Rybnica</i> , Sln. <i>Ribnica</i>	toponym

Examples of PIE compounding persist in relic form, e.g., OCS *gospodь* ‘the Lord’, ‘lord’ (< PIE \*g<sup>h</sup>osti- + \*pot-, if it is in fact cognate with Lat. *hospes*, *hospitis* – Slavic *d* is unexpected as a reflex of PIE \*t), or the famous taboo word for ‘bear’, OCS *medvěďb* (< PIE \*med<sup>h</sup>u- ‘honey’ + \*ed- ‘eat’), cf. *o*-grade USorb. *sknadž*, Sln. *strnad*, Ukr. *strenadka* ‘yellowhammer (*Emberiza citrinella*)’ (< CSL. \*strъnadь < PIE \*kri(m)n(o)- ‘grain’ + \*h<sub>1</sub>od- ‘eat(er)’) (Snoj 1992: 198–199). Slavic given names preserve further endocentric compounds that were productive in the medieval pre-Christian era, e.g., ORuss. *Svätoslavь* (← PIE \*kwen-to- ‘sacred, strong, good’ + \*kle/ow- ‘famed’), OCS *Vladimirь* (← PIE \*wel-d<sup>h</sup>e- ‘rule, have strength’ + \*meyH-ro- ‘good, pleasing’).

## SYNTAX

The syntax of Slavic languages is treated less in the historical than in the synchronic spectrum of Slavic linguistics than are phonology and morphophonemics, which have received the majority of attention. In contrast, however, studies of Slavic synchronic syntax have burgeoned in the last several decades. Here we will focus on some highlights with a particular focus on developments relating the Indo-European heritage of Slavic.

Generally, Slavic simple clause structure tends to be SVO, though word order is “free,” presumably owing to case marking. Much attention in the structural literature, particularly the Prague School, has focused on the functional sentence perspective in Slavic languages. So, for example, in contemporary Bulgarian (a language that has reduced case distinctions down to three) one may produce a natural three-element sentence such as SVO *Ivan nameri knigata* ‘Ivan found book-the’ also as VSO, VOS, SOV, OVS, and OSV, all responding to the questions *Kakvo stana?* ‘What happened?’ or *Kakvo pravi Ivan?* ‘What is Ivan doing’, where the focus of the sentence changes depending on the sentence intonation, the element in question being given emphatic intonation (examples from Kempgen et al. 2009: 656–657). A traditional Prague School analysis (without regard to sentence intonation) assumes that the focus (“new information”) is the last sentence element, a pattern that generally holds in written Slavic languages.

From an Indo-European perspective one may note the persistence of Wackernagel’s Law in OCS and modern Slavic languages, which assigns unstressed particles to the

second position in the clause following the first accented element *oni že vložišę rǫčę na nę i ęsę ję* ‘they [focus particle] placed hands on him and took him’ (Codex Marianus, Mark 14:46), cf. *Cze zdá se mně* ‘it seems to me’, Sln. *zdi se mi* ‘idem’. East Slavic is exceptional in that it has fused the reflexive particle *šja* (< CSI. \*sę) to the right of the verb (e.g., Russ. *kusat* ‘to bite’ vs. *kusat’sja* ‘to bite one another, be characterized by biting’), and the trend toward fusion can be seen in the chronological progression of texts in the Novgorod Birchbark corpus: *a nýne sã družina po mã poručila* (no. 109, late 12th c.) ‘and now [REFL.] retinue about me vouched’, *a jazo sã klaneju* (no. 344, early 14th c.) ‘and I REFL. bow’ (salutation formula), *a čemu sã gněvaeši?* (no. 605, 12th c.) ‘but why REFL. anger-[2 sg.]’ vs. *a ázo tobe klanãjusã* (no. 147, 13th c.) ‘and I to-you bow-REFL.’, *a to sã dijalošb sedně vo Veliki dnę* ‘and this REFL. done-REFL. today on Great (Easter) day’ (no. 154, 15th c.), with the latter example reflecting a rare simultaneous instantiation of both the older and clitic placement (Zaliznjak 2004: 188–189).

Significant variation is found in compound sentences in the Slavic languages, which underwent major shifts in the historical period. Canonical OCS attests to a strategy of clause combining employing participles in the dative case for subordination (the “dative absolute”), which subsequently disappeared from most of Slavic save for southern Russian dialects (see Andersen 1970), e.g., OCS *vъ pętoe že na desęte lęto vladýčystva tiveręja k’esarja obladajořtu pontęskumu pilatu ijudęeję, i četvęrtovlastęvųjořtu galileeję irodu . . . bystę glagolę božii kę ioanu zaxarinu synu* ‘in fifth-[def.] year of reign Tiberius-[GEN.] Caesar-[GEN.] ruling-[PRES. ACT. PTCP. DAT. SG. M.] Pontius-[DAT.] Pilate-[DAT.] and ruling-as-tetrarch-[PRES. ACT. PTCP. DAT. SG. M.] Galilee Herod-[DAT.] . . . was-[AOR. 3 SG.] word-[NOM. SG. M.] God’s to John Zakharias’ son – ‘In the fifteenth year of the reign of Tiberius Caesar, Pontius Pilate being governor of Judea, and Herod being tetrarch of Galilee, . . . the Word of God came unto John the son of Zacharias’ (Luke 3:1–2, Codex Zographensis); *тѣного сѣшту народу i ne imořtemę česo ěsti . . . isusę glagola* ‘being-[PRES. ACT. PTCP. -DAT. PL] many people and [they] not having-[PRES. ACT. PTCP. -DAT. PL] what-[GEN. SG.] to-eat . . . Jesus spoke’ (Mark 8:1, Vaillant 1950–1977: 90). Variation in the material of the general (realis) subordinating conjunctions in Slavic languages reflects the degree to which subordination has been reshaped in Slavic languages, e.g., East Slavic *čto* (< CSI. \*čto ‘what’ < PIE \*k<sup>w</sup>i- + \*to-), West Slavic *že* (< CSI. \*že ‘focus marker’ < PIE \*g<sup>wh</sup>e), Sln., BCMS *da* (< PIE \*doh<sub>2</sub> ‘lative particle’, cf. Germ. *zu*, Eng. *to*), Prekmurje Sln., Međimurje Kajkavian *ka* (< \*k<sup>wh</sup>eh<sub>2</sub>, cf. Lat. *quā* ‘in what manner’), Mac. *deka* (< PIE \*d<sup>h</sup>e ‘there’ + \*k<sup>wh</sup>eh<sub>2</sub>), Bulg. *če* (< PSI. \*ke < PIE \*k<sup>w</sup>e ‘and’) and *deto*. A similar pattern of variation occurs with relative markers. OCS mainly employed a compound of \*j- pronouns + the *že* marker (e.g., *xlębъ bo jęžę azъ damъ plętъ moę estę jęžę azъ damъ za životę vъsego mira* ‘bread thus REL.-ACC.-SG.-M. I give-1-sg. flesh my-f.-sg.-nom. is-2-sg. REL.-ACC.-SG.-F. I give-1-sg. for life whole-gen.-sg.-m. world’; John 6:51, Codex Zographensis). Already in the Freising Folia an emerging phonological change *ž* > *r*, which spread throughout South Slavic, reshaped the marker, setting up the conditions by which the inherited relative adjective \*kęter- (< PIE \*k<sup>w</sup>otero- ‘which of two’) could be reanalyzed as *kęte-* + *r*, where the last element became a generalized relativizer. The *-r* element subsequently became a productive formant for building a range of relativizers in Sln. and dialects of BCMS associated with the Roman rite (Kajkavian, Čakavian) (e.g., Sln. *kar* ‘that which’, *kadar* ‘when[ever]’, *kdor* ‘who[ever]’) and was replaced in dialects associated with the Byzantine rite (see Greenberg 1999 for details).

A further step in clause combining is the loss of the infinitive in the Balkan Slavic languages, Mac. and Bulg., an areal feature characteristic of the Balkan languages (Albanian,

Aromanian, Greek, Romanian) that has also spread to Serbian and Montenegrin. This feature, as well as others discussed in this chapter, may be observed in the following fragment from Jaroslav Hašek's *Osudy dobrého vojáka Švejka za světové války* (The adventures of the good soldier Schweik during the world war) in the original Czech and several translations. The fragment translates “‘Which Ferdinand [did they kill], Mrs. Müller?’” asked Schweik, not ceasing to massage his knees, ‘I know two Ferdinands.’”

### West Slavic

- *Którego [1] Ferdynanda, pani Müllerowa? – zapytał [2] Szwejk, nie przestając [3] masować [4] kolan. – Ja znam dwóch Ferdynandów.* (Polish)
- *“Kterýho Ferdinanda, paní Müllerová?” otázal se Švejk, nepřestávaje si masírovat kolena, “já znám dva Ferdinandy.”* (Czech)
- *“Ktorého Ferdinanda, pani Müllerová?” opýtal sa Švejk, neprestávajúc si masírovat kolená. “Ja poznám dvoch Ferdinandov.”* (Slovak)

### Western South Slavic

- *“Katerega Ferdinanda, gospa Müllerjeva?” je vprašal Švejk, ne da bi si nehal mazati kolena, “poznam dva Ferdinanda [5].”* (Slovene)
- *Kojega Ferdinanda, gospođa Müllerova? – odazvao se Švejk ne prestajući masirati koljena — ja poznam dva Ferdinanda.* (Croatian)
- *Koga to Ferdinanda, gospođo Miler? – upita Švejk, ne prekidajući masiranje kolena.- Ja poznajem dva Ferdinanda.* (Serbian)

### Eastern (Balkan) South Slavic

- *Koj Ferdinand, gospođo Miler? – zapraša Švejk, ne prestanuvajki da si gi masira kolenata [6], – poznavam dva Ferdinanda.* (Macedonian)
- *Koj Ferdinand, gospožo Mjulerova? – zapita Švejk, bez da prekъsne da raztirava kolēnite si. – Az [7] znaja dvama Ferdinandovci.* (Bulgarian)

### East Slavic

- *Jakoho ce Ferdinanda, pani Mjullerova? – spytav Švejk, ne perestajučy roztyraty kolina. – Ja znaju dvox Ferdinanda.* (Ukrainian)
- *Jakoha Ferdynanda, Pani Mjuler? – spytav Švejk, ne perastajučy masirovac 'kaleni. – Ja vedaju dvux Ferdynandaw.* (Belarusian)
- *Kakogo Ferdinanda, pani Mjullerova? – sprosil Švejk, ne perestavaja massirovat' koleni. – Ja znaju dvux Ferdinandov.* (Russian)

### Non-Slavic Balkan

- *Pe care Ferdinand, doamnă Müllerová? întrebă Švejk continuind să-și maseze genunchiul. Eu cunosc doi.* (Romanian)
- *Cilin Ferdinand, zonja Myler? – pyeti Shvejku duke fërkuar gjunjët. – Unë njoh dy Ferdinandë.* (Albanian)

- 1 With the exception of Mac., Bulg., which no longer mark cases in nouns and adjectives, the M. SG. ACC. animate forms in Slavic take the marking of the genitive.
- 2 Past tense ‘he asked’ is rendered with the aorist (simple past) 3 SG. in Mac., Bulg., which preserve both simple and coordinated past tenses, and the relic of the perfect in the rest of the Slavic languages, which have only one past tense construction. The simple past 3 SG. is used in Romanian (*întrebă*) and Albanian (*pyeti*). Sln. preserves the full construction with clitic auxiliary *je* ‘be’ 3 SG. (cf. OCS *jestъ*) + the *l*-participle. Auxiliary clitics are obligatory in non-3-PERS. PAST tense in all but the East Slavic languages, which use the unauxiliated *l*-participle for all persons and numbers. The 3 SG. auxiliary has been eliminated in West Slavic (though it persisted well into the historical period), while in BCMS it is normally dropped in reflexive constructions (*odazvao se* ‘he responded’, but would have been *pitao je* ‘he asked’, had the translator chosen to use this verb).
- 3 With the exception of Sln. and Bulg. the first verb in the subordinate clause, ‘not stopping’, is rendered with a PRES. ACT. ADV. PTCP. (PSl. \*-ANT-Y-). The option would have been available to both of these languages as well but was not chosen by the translator, though in Sln. the participle is deprecated with all but a few lexical items.
- 4 Except for Mac., Bulg., the Slavic translations use the inf. form of ‘to massage’. In Mac. the equivalent construction contains *da* COMP., *si* REFL.-DAT. (i.e., ‘one’s own’), *gi* them.-ACC.-PL. (agreeing with an object marked for definiteness), *masira* ‘massage’ 3-SG.-IND. In this regard, Mac. takes the infinitival phrase as the “Wackernagel” unit, placing all of the clitics after the participle. In Bulg. the construction is similar, though the possessive marker *si* is moved to the right of the object *kolenite si* ‘knees-the REFL.-DAT.’ Romanian is similar to the Mac. construction, though it lacks an equivalent to the reduplicated *gi* definite-object marker. Albanian has developed an innovative infinitival construction *duke fërkuar*, though a finite verb would also have been possible, *pa ndaluar së fërkuari gjunjët* ‘without continuing subordinating conjunction rub-3-SG. knees’ (see chapter on Albanian).
- 5 Sln. retains a distinct dual form, which also corresponds to dual markers in nominal and verbal inflection. In East Slavic the dual has expanded to include groups of three and four (“paucals”), marked only in nominal phrases.
- 6 Mac. and Bulg. mark definiteness with the form of the inherited DEM. PRON. CSI. \*тЪ (agreeing in gender and number with its head), marked to the right of the first element of the nominal phrase. Parallel constructions are found in Romanian and Albanian.
- 7 Sln., Mac., and Bulg. continue the full form of the 1 SG. pronoun (*jaz*, *jas*, *az*, respectively), though in PRO-drop languages (West and South Slavic) the personal pronouns are not obligatory except for discursive strategies such as reference switching and emphasis.

## FURTHER READING

Students of Slavic languages can get a good overview of Slavic from handbooks, including the English-language reference work Comrie and Corbett 1993, which is arranged according to standard languages; a more concise Russian-language work, Moldovan 2005, is structured similarly. An English and German handbook, Kempgen et al. 2009, is arranged thematically but also includes sketches of standard languages. A single-volume overview of Slavic philology suitable for advanced undergraduate students is Schenker

1995. Several English-language grammars of Slavic languages, written by specialists and intended for general linguistic typologists and comparativists, are available freely on the Internet on the SEELRC Network (<http://www.seelrc.org/projects/grammars.ptml>). An annotated bibliography of the literature up to the early 1980s on the reconstruction of Common Slavic is available in Birnbaum 1979 and Birnbaum and Merrill 1983. The long-term multinational project on Slavic diatopy, the Common Slavic Linguistic Atlas, published in various places in book format, is now available in electronic format (see the *Obščeslavjanskij lingvističeskij atlas*). The etymology of Slavic languages as a whole is covered in the as yet unfinished series begun under the editorship of the late O. N. Trubačev (1974–); Kopečný 1973, 1980 remain an invaluable source for the etymology of pronouns, particles, and grammatical elements. Single-volume etymological dictionaries that take into account recent research include Gluhak 1993 (which traces words back to Nostratic), Snoj 2003, and Derksen 2008. Kopečný 1981 provides a single-volume listing of Slavic comparanda for the ca. 2000 words of CSI. core vocabulary. Among the historical grammars of Slavic one could view as classic the works of Arumaa (1964–1985) and Vaillant (1950–1977), the latter remaining the most comprehensive treatment of Slavic comparative grammar to date. With regard to historical phonology, although dated in some respects (particularly with regard to the development of Slavic word prosody), Shevelov 1964 remains extremely valuable for the sheer amount of documentation for Slavic in a single volume. Shevelov also founded a series of (English-language) historical phonologies of Slavic languages that is still in progress, of which volumes on Belarusian, Macedonian, Polish, Slovak, Slovene, Sorbian, and Ukrainian have appeared to date. Of special note for the reconstruction of Slavic accent would be the watershed works of Stang 1957 and Illič-Svityč 1963, the latter of which is also available in English translation (1979). These works have renewed energy in research on Baltic and Slavic accentology with particular regard to its Indo-European origins, notable monographs of which range from Kortlandt 1975, Garde 1976, Dybo 1981 and 2000 to the works of the newest generation of scholars, e.g., Olander 2009, Pronk 2012, and Kapović 2015. Numerous handbooks of OCS exist in Slavic and other European languages, most of which are to a greater or lesser extent derivative of the classic textbook by Leskien (1969 [1871]); Lunt 2001 is currently the best OCS reference in English. The authoritative OCS dictionary is the multi-volume Kurz et al. 1958–1997, while Sadnik and Aitzetmüller 1955 and Cejtin et al. 1994 are one-volume works. Sadnik and Aitzetmüller also contains a reverse dictionary of OCS words. Outside of canonical OCS, medieval attestations of Slavic include the Novgorod Birchbark Letters, which are available mostly in Russian sources, of which Zaliznjak 2004 gives the most comprehensive overview; and the Freising Folia, which attests to a variety of Slavic spoken around 1000 A.D. and is most closely associated with Slovene (see Bernik 1993, which also contains English glosses and translation). A recently published work (Boček 2014) gives an overview of competing theories of the development of Proto-Slavic, in particular the notion of Late Common Slavic as a *lingua franca*, as well as a discussion of frameworks for language contact with Slavic.

## NOTES

- 1 A similar stylization of the phonological system of the reconstructed systems is employed by Holzer, whose system corresponds to both the PSI. and CSI. systems in this chapter. The details of Holzer's system are found in Holzer 2007, esp. pp. 15–20.
- 2 Possibly also from \*ǵ<sup>h</sup>woy-st-, cf. OPruss. *svāigstan* 'light, shining' (see Gluhak 1993, s.v.).



## REFERENCES

- Andersen, Henning 1969, "Lenition in Common Slavic," *Language* 45: 553–574
- Andersen, Henning 1970, "The Dative of subordination in Baltic and Slavic," in: Magner, Thomas F. & Schmalstieg, William R. (eds.), *Baltic Linguistics*, Pennsylvania State University Press, University Park–London: 1–9
- Andersen, Henning 1996, *Reconstructing Prehistorical Dialects. Initial Vowels in Slavic and Baltic (Trends in linguistics, studies and monographs 91)*, Mouton de Gruyter, Berlin–New York
- Andersen, Henning 1999, "The western South Slavic contrast Sn. *sah-ni-ti* || SC *sah-nu-ti*," *Slovenski jezik / Slovene Linguistic Studies* 2: 47–62
- Andersen, Henning 2003, "Slavic and the Indo-European migrations," in: Andersen, Henning (ed.), *Language Contacts in Prehistory: Studies in Stratigraphy (Current Issues in Linguistic Theory 239)*, Benjamins, Amsterdam: 45–76
- Andersen, Henning 2013, "On the Origin of the Slavic Aspects: Aorist and Imperfect," *Journal of Slavic Linguistics* 21/1: 17–43
- Arumaa, Peeter 1964–1985, *Urslavische Grammatik I–III*, Winter, Heidelberg
- Bernik, France (ed.) 1993, *Brižinski spomeniki. Znanstvenokritična izdaja (SAZU, Razred za filološke in literarne vede 39)*, ZRC SAZU, Ljubljana
- Birnbaum, Henrik 1979, *Common Slavic: Progress and Problems in its Reconstruction*, Slavica, Columbus
- Birnbaum, Henrik & Merrill, Peter T. 1983, *Recent Advances in the Reconstruction of Common Slavic (1971–1982)*, Slavica, Columbus
- Boček, Vít 2014, *Praslovanština a jazykový kontakt*, Lidové noviny, Prague
- Cejtlin, R. M.; Večerka, R. & È. Blagovaja (eds.) 1994, *Staroslavjanskij slovar' (po rukopisjam X–XI vekov)*, Russkij jazyk, Moscow
- Comrie, Bernard & Corbett, Greville C. (eds.) 1993, *The Slavonic Languages*, Routledge, London–New York
- Curta, Florin 2004, "The Slavic lingua franca (Linguistic notes of an archeologist turned historian)," *East Central Europe* 31/1: 125–148
- Darden, Bill J. 1990, "On the history and function of ablaut in the verb from Balto-Slavic to Russian," *University of Chicago Working Papers in Linguistics* 5: 1–15
- Derksen, Rick 2008, *Etymological Dictionary of the Slavic Inherited Lexicon*, Brill, Leiden–Boston
- Dickey, Stephen M. 2000, *Parameters of Slavic Aspect: A Cognitive Approach*, CSLI Publications, Stanford
- Dickey, Stephen M. 2005, "S-/Z- and the grammaticalization of Slavic aspect," *Slovenski jezik / Slovene Linguistic Studies* 5: 3–55
- Dickey, Stephen M. 2011, "The varying role of *po-* in the grammaticalization of Slavic aspectual systems: Sequences of events, delimitatives, and German language contact," *Journal of Slavic Linguistics* 19/2: 175–230
- Dickey, Stephen M. Forthcoming, *The Evolution of Slavic Verbal Aspect*, Mouton de Gruyter, Berlin–New York
- Doluxanov, Pavel M. 2000, *Istoki ètnosa, Evropejskij dom*, Saint Petersburg
- Dybo, Vladimir Antonovič 1981, *Slavjanskaja akcentologija. Opyt rekonstrukcii sistemy akcentnyx paradim v praslavjanskom*, Nauka, Moscow
- Dybo, Vladimir Antonovič 2000, *Morfonologizovannye paradigmatičeskie akcentnye sistemy. Tipologija i genezis*, tom 1, Jazyki ruskoj kul'tury, Moscow

- Gamkrelidze, Tamaz & Ivanov, Vjačeslav Vs. 1973, "Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse," *Phonetica* 27: 150–156
- Garde, Paul 1976, *Histoire de l'Accentuation slave*, Tome I–II, Collection de manuels de l'Institut d'études slaves, Paris
- Gluhak, Alemko 1993, *Hrvatski etimološki rječnik*, August Cesarec, Zagreb
- Greenberg, Marc L. 1999, "Multiple causation in the spread and reversal of a sound change: Rhotacism in South Slavic," *Slovenski jezik / Slovene Linguistic Studies*: 63–76
- Greenberg, Marc L. 2007, "Phonetic evidence for the development of the 'acute' tone in Slavic," in: Kapović, Mate & Matasović, Ranko (eds.), *Tones and Theories: Proceedings from the International Workshop on Balto-Slavic Accentuation*, Institut za hrvatski jezik i jezikoslovlje, Zagreb: 95–108
- Greenberg, Marc L. 2010, "PIE inheritance and word-formational innovation in Slavic motion verbs in *-i-*," in: Driagina-Hasko, Viktorija & Perelmutter, Renee (eds.), *New Approaches to Slavic Verbs of Motion*, Benjamins, Amsterdam: 111–121
- Greenberg, Marc L. 2011, "A Balkanism in Central Europe? Realis vs. irrealis in subordinate clauses in Prekmurje Slovene," in: Holub, Zbyněk & Sukač, Roman (eds.), *Dialektologie a geolingvistika v současné střední Evropě* Slezská univerzita, Opava: 8–18
- Greenberg, Marc L. & Dickey, Stephen M. 2006, "Slavic \*jazditi 'to ride' and its implications for the development the category of (in)determinacy," in: Jesenšek, Marko & Zorko, Zinka (eds.), *Jezikovna predanost. Akademiku prof. dr. Jožetu Toporišiču ob 80-letnici (Zbirka Zora 44)*, Slavistično društvo–SAZU, Maribor–Ljubljana: 153–158
- Holub, Zbyněk 2014, *Jazykové doklady prehistorické podunajské migrace slovanského etnika na český jih*, Slezská univerzita/OPTYs, Opava
- Holub, Zbyněk & Greenberg, Marc L. 2013, "A circum-Pannonian word-prosodic parallel: paroxytonic accent in the south-west Bohemian dialect," *Jezikoslovni zapiski* 19/2: 11–26
- Holzer, Georg 2007, *Historische Grammatik des Kroatischen. Einleitung und Lautgeschichte der Standardsprache (Schriften über Sprachen und Texte 9)*, Peter Lang, Frankfurt
- Illič-Svityč, Vladislav Markovič 1963, *Imennaja akcentuacija v baltijskom i slavjanskom: sud'ba akcentuacionnyx paradigm*, Nauka, Moscow [English translation: Feldstein, Ronald (trans.) 1979, *Nominal accentuation in Baltic and Slavic*, The MIT Press, Cambridge–London]
- Ivanov, Vjačeslav Vs. 1981, *Slavjanskij, baltijskij i rannebalkanskij glagol. Indoevropskie istoki*, Nauka, Moscow
- Jakobson, Roman 1963, "Opyt fonologičeskogo podxoda k istoričeskim voprosam slavjanskoj akcentologii. Pozdnij period slavjanskoj jazykovoj praistorii," in: *American Contributions to the Fifth International Congress of Slavists, Sofia, September 1963*, vol. 1: *Linguistic Contributions*, Mouton, The Hague: 153–178
- Janda, Laura 1996, *Back from the Brink: A Study of how Relic forms in Languages Serve as Source Material for Analogical Extension (LINCOS Studies in Slavic Linguistics 1)*, LINCOS Europa, Munich
- Kapović, Mate 2009, "The Accent of Slavic \*ja(зъ) 'I'," in: Olander, Thomas & Larsson, Jenny H. (eds.), *Stressing the Past: Papers on Baltic and Slavic Accentology (Studies in Slavic and General Linguistics 35)*, Rodopi, Amsterdam–New York: 53–73
- Kapović, Mate 2015, *Povijest hrvatske akcentuacije. Fonetika*, Matica hrvatska, Zagreb

- Kempgen, Sebastian, Peter Kosta, Tilman Berger & Karl Gutschmidt (eds.) 2009, *Die Slavischen Sprachen: Ein Internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung / The Slavic Languages: An International Handbook of their Structure, their History and their Investigation*, vol. 1 (*Handbooks of Linguistics and Communication Science* 32.1), de Gruyter, Berlin–New York
- Kopečný, František 1973, 1980, *Etymologický slovník slovanských jazyků. Slova gramatická a zájmena. I Předložky, koncové partikule; II Spojky, částice, zájmena a zájmenná adverbia*, Academia, Prague
- Kopečný, František 1981, *Základní všeslovanská slovní zásoba*, Academia, Prague
- Kortlandt, Frederik 1975, *Slavic Accentuation*, de Ridder, Lisse
- Kortlandt, Frederik 1978, "Proto-Indo-European obstruents," *Indogermanische Forschungen* 83: 107–118
- Kortlandt, Frederik 1985, "Proto-Indo-European glottalic stops: The comparative evidence," *Folia linguistica historica* 6/2: 183–201
- Kurz, Josef, et al. 1958–1997, *Slovník jazyka staroslovanského*, ČSAV, Prague
- Lehfeldt, Werner 2001, *Einführung in die morphologische Konzeption der slavischen Akzentologie, 2., verbesserte und ergänzte Auflage mit einem Appendix von Willem Vermeer (Vorträge und Abhandlungen zur Slavistik, Bd. 42)*, Otto Sagner, Munich
- Leskien, August 1969 (1871), *Handbuch der altpulgarischen (altkirchenslavischen) Sprache<sup>9</sup>*, Winter, Heidelberg
- Lunt, Horace G. 1984–85, "On Common Slavic," *Zbornik Matice srpske za filologiju i lingvistiku* XXVII–XXVIII: 417–422
- Lunt, Horace G. 2001, *Old Church Slavonic Grammar<sup>7</sup>*, de Gruyter, Berlin–New York
- Miller, Raymond H. 1988, "The third person present tense and Common Slavic dialectology," *International Journal of Slavic Linguistics and Poetics* XXXVII: 7–33
- Moldovan, Aleksandr Mixailovič (ed.) 2005, *Slavjanskije jazyki (Jazyki mira)*, Nauka, Moscow
- Nichols, Johanna 1993, "The linguistic geography of the Slavic expansion," in: Timberlake, Alan (ed.), *American Contributions to the Eleventh International Congress of Slavists*, Slavica, Columbus: 377–391
- Obščeslavjanskij lingvističeskij atlas* <http://www.slavatlas.org/>
- Olander, Thomas 2009, *Balto-Slavic Accentual Mobility (Trends in Linguistics, Studies and Monographs 199)*, de Gruyter, Berlin–New York
- Pleterski, Andrej 2013, *The Invisible Slavs. Župa Bled in the "prehistoric" Early Middle Ages (Opera Instituti Archaeologici Sloveniae 28)*, Inštitut za arheologijo ZRC SAZU, Ljubljana
- Pronk, Tijmen 2012, "Proto-Indo-European long vowels and Balto-Slavic accentuation," *Baltistica* XLVII (2): 205–247
- Sadnik, Linda & Aitzetmüller, Rudolf 1955, *Handwörterbuch zu den altkirchenslavischen Texten*, Mouton, s'Gravenhage
- Schallert, Joseph & Greenberg, Marc L. 2007, "The prehistory and areal distribution of Slavic \*gьlčěti 'speak'," *Slovenski jezik / Slovene Linguistic Studies* 6: 9–76
- Schenker, Alexander M. 1995, *The Dawn of Slavic: An Introduction to Slavic Philology*, Yale University Press, New Haven
- Shevelov, George Y. 1964, *A Prehistory of Slavic. The Historical Phonology of Common Slavic*, Winter, Heidelberg
- Sihler, Andrew L. 1995, *New Comparative Grammar of Greek and Latin*, Oxford University Press, New York–Oxford

- Snoj, Marko 1992, "K preučevanju starejših nominalnih kompozit: psl. \**drozdъ* in \**strъnadъ*," *Miklošičev zbornik (Obdobja 13)*, SAZU–Univerza v Ljubljani, Ljubljana: 193–200
- Snoj, Marko 2003, *Slovenski etimološki slovar*<sup>2</sup>, Modrijan, Ljubljana
- Snoj, Marko & Greenberg, Marc L. 2012, "O jeziku slovanskih prebivalcev med Donavo in Jadranom v srednjem veku (pogled jezikoslovcev) / On the language of the medieval Slavic population in the area between the Danube and the Adriatic (from a linguistic perspective)," *Zgodovinski časopis / Historical Review* 66/3–4 (146): 276–305
- Stang, Christian S. 1957, *Slavonic accentuation*, Nygaard, Oslo
- Trubačev, Oleg Nikolaevič 1968, *Nazvanija rek pravoberežnoj Ukrainy. Slovoobrazovanie. Ètimologija. Ètničeskaja interpretacija*, Nauka, Moscow
- Trubačev, Oleg Nikolaevič 1974–, *Ètimologičeskij slovar 'slavjanskix jazykov. Praslavjanskij leksičeskij fond*, 1: (A—\**besědъlivъ*)—, Nauka, Moscow
- Vaillant, André 1950–1977, *Grammaire comparée des langues slaves I–V*, IAC–Klincksieck, Lyon–Paris
- Vermeer, Willem 1984, "On clarifying some points of Slavonic accentology: The quantity of the thematic vowel in the present tense and related issues," *Folia linguistica historica* 5/2: 331–395
- Vermeer, Willem 1994, "Why the early North Russian nominative singular in -e does not palatalize stem-final velars," *Russian Linguistics* 18/2: 145–157
- Vermeer, Willem 2014, "Early Slavic dialect differences involving the consonant system," *Dutch Contributions to the Fifteenth International Congress of Slavists, Minsk: Linguistics (Studies in Slavic and General Linguistics 40)*, Rodopi, New York: 181–227
- Winter, Werner 1978, "The distribution of short and long vowels in stems of the type Lithuanian *èsti* : *vèsti* : *mèsti* and OCS *jasti* : *vesti* : *mesti*", in: Fisiak, Jacek (ed.), *Recent Developments in Historical Phonology (Trends in Linguistics. Studies and Monographs 4)*, de Gruyter, The Hague–Paris–New York: 434–445
- Zaliznjak, Andrej Anatol'evič 2004, *Drevnenovgorodskij dialekt*<sup>2</sup>, Jazyki slavjanskoj kul'tury, Moscow

# ALBANIAN

*Alexander Rusakov*

## INTRODUCTION

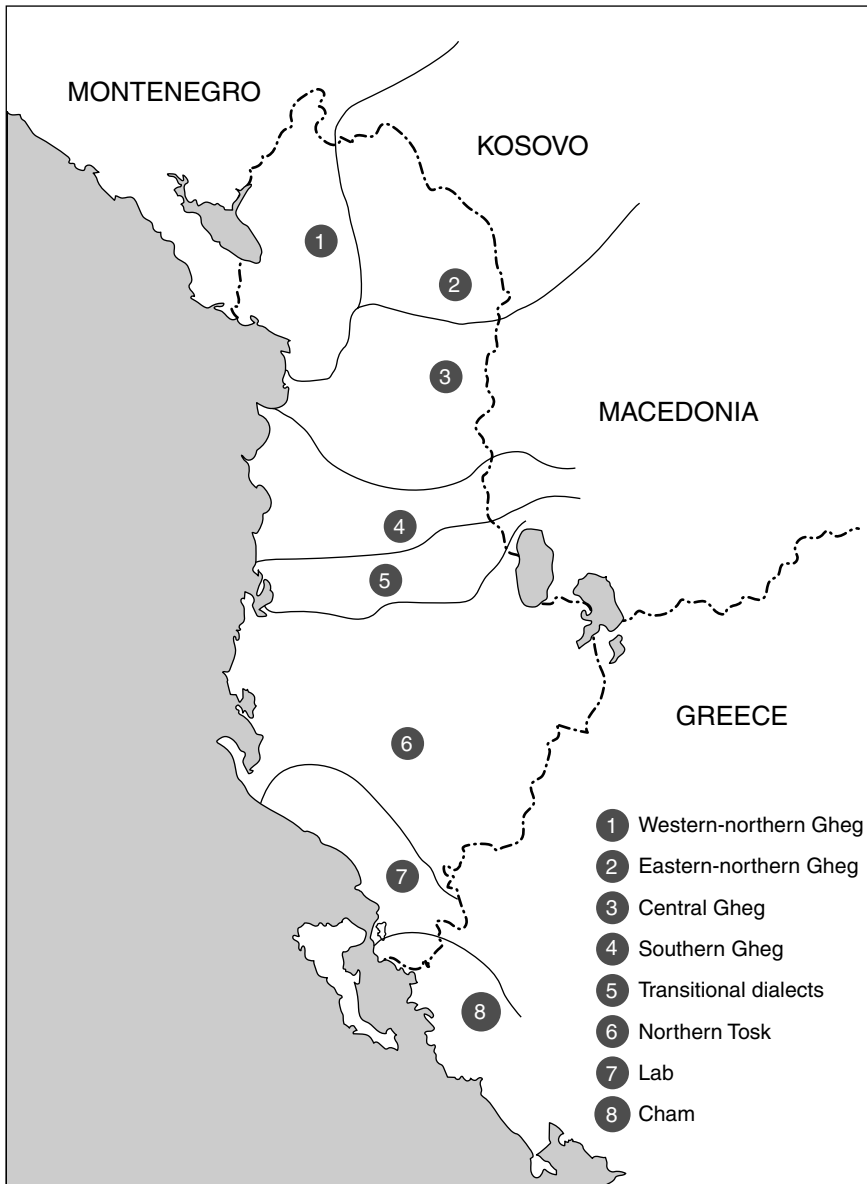
*Current geographic distribution.* Albanian is currently spoken in Albania, Kosovo, the southern part of Montenegro, the southern part of Serbia, the western part of Macedonia and, to some extent, the northwestern part of Greece. Sizeable old Albanian diasporas exist in various regions of Greece (since the late 13th–14th centuries), in the southern Italy (since the 15th century), near Zadar in Croatia (since the 18th century), in Bulgaria and Ukraine (this population probably left Albania in the 16th century). Albanian is spoken by approximately 7.5 million speakers.

*Dialects.* In terms of dialectal variation, the Albanian language area is divided into two large zones – the Northern (or *Gheg*) and the Southern (or *Tosk*); the Shkumbin river, which crosses Albania from east to west, forms a border between the two zones (a narrow strip of transitional dialects is situated to the south of the Shkumbin river). The dialects of northern Albania, Kosovo, Montenegro, and Serbia; the diaspora dialect of Zadar; and the majority of the Albanian dialects of Macedonia belong to the Gheg zone, while all other dialects belong to the Tosk zone.

The main differences between the Gheg and Tosk dialects in phonetics are (1) the so-called rhotacism (the change *-n-* > *-r-* in Tosk); (2) the presence of nasal vowel series in Gheg and the absence of nasal vowels in Tosk; (3) the correspondence between Gheg stressed nasal *ã* and Tosk stressed *ə*; (4) the correspondence between Gheg *vo-* and Tosk *va-* in some words. Concerning all these phenomena, the Gheg dialect features a more archaic language state than the Tosk dialect. Other phonetic isoglosses are not absolute. On the basis of the distribution of these phonetic phenomena in lexical borrowings one may conclude that the division of (pre)Albanian language continuum into the Gheg and Tosk zones had to occur in the last centuries of the first millennium A.D.

The main morphological difference between the Gheg and Tosk dialects, which is also almost absolute, is the presence of the analytical infinitive form in Gheg. The Tosk dialect lacks the infinitive form *strictu sensu* (except some lexicalized remnants). There are some other less important morphological differences between the two dialects (e.g. lack of a possessive-reflexive pronoun in Tosk, some differences in word-formation affixes, and so on). Some other phonetic and morphological differences will be listed in the relevant sections of this chapter.

Both the Gheg and Tosk zones are divided into numerous subdialects. The main Gheg subdialects are northern Gheg (further subdividing into western and eastern subdialects), central Gheg and southern Gheg (or Gheg of Central Albania). The Tosk dialect divides into northern Tosk (further subdividing into western and eastern subdialects) and southern Tosk (Lab and Cham dialects). The main diaspora dialects, *Arbresh* (the Albanian dialects of Italy) and *Arvanitika* (the Albanian dialects of Greece), are closest to southern Tosk, although they (especially Arbresh) also demonstrate some specific dialectal traits due to the complex scenarios of their respective speakers' migrations. Both dialects are very important for Albanian language history.



**MAP 12.1** THE ALBANIAN DIALECTS

Source: Adapted from: Gjinari et al. 2007, map C and Gjinari J., Shkurtaj Gj. 2009, *Dialektologji*, Shtëpia botuese e librit universitar, Tiranë: p. 161

Modern Standard Albanian was consolidated after the Second World War on the basis of the Tosk dialect but including some elements (mainly, lexical and, to a lesser degree, (morpho)phonological and morphological) of the Gheg dialect. Now the standard language is used as such in all parts of the Albanian language area.

*Documentation of Albanian.* The first well-documented mention of Albanians as a separate ethnic entity is encountered in the “Historia” by Michael Attaleiates (1079–1080), where Albanians are mentioned with respect to the Durrës region in the year 1078 (two earlier mentions under the years 1038 and 1043 are disputable). The first known mention of Albanian as a separate language is found in a legal document from a Raguza (Dubrovnik) archive dated 1285: *Audivi unam vocem clamantem in monte in lingua Albanesca* ‘I heard a voice on the mountain that shouted in the Albanian language’. The first Albanian language record (in the Gheg dialect) is the baptismal formula (*Unte paghesont premenit Atit et birit et spertit senit* ‘I baptize you in the name of the Father and the Son and the Holy Ghost’) from 1462, included in the Latin pastoral letter of Archbishop of Durrës Pal Engjëlli (Paulus Angelus), a close associate of Scanderbeg. There are some other short language records (in both Gheg and Tosk) belonging to the end of 15th and beginning of the 16th centuries (see Roques 1932, Shuteriqi 1976, Matzinger 2010).

The first important early Albanian written text is “*Meshari*” (Missal) by the Catholic priest *Gjon Buzuku*, printed in 1555 (most likely in Venice); only one copy is preserved, and the text comprises 188 pages. “*Meshari*” was written in the northern variety of the Gheg dialect; it remains the main source for the study of the earliest stage of Albanian. “*Meshari*” opens the *North Albanian Catholic tradition* in the development of Albanian writing. Overall, some ten printed texts belonging to this tradition and published before the middle of 18th century are known. All these publications use Latin script with rather similar graphic and orthographic conventions. These books are mostly devoted to spiritual and church matters (however, there was also a *Dictionarium latino-epiroticum* (Roma 1635) by *Frang Bardhi* (Franciscus Blanchus)). Besides Catholic literature, some other traditions in the early Albanian writing can be identified. One of these traditions is the *Orthodox tradition of Central and South Albania*, which is represented by the texts created between the second third of the 18th and the first third of the 19th centuries. These texts were written in southern Gheg and Tosk and used both Greek script and various original alphabets invented by educated Albanians especially for their own language needs. Along with a small number of religious texts, two multilingual dictionaries of great value belong to this tradition: *Theodor Kavalioti’s “Prôtopeiria”* (1770) and *Daniel of Voskopojaja’s “Eisagôgikê didaskalia”* (1802). The *Muslim tradition of Old Albanian writings* covered the entire territory of Albania but was especially prosperous in the cities of Central Albania, where some poets of a high poetical level wrote their verses using Arabic script from the beginning of the 18th till the middle of the 19th century. Great significance for the study of Albanian belongs to the *Italo-Albanian (Arbëresh) writing tradition*, which started with “*E mbsuame e krështere*” (Christian Doctrine) by *Luca Matranga* (published in 1592 in Rome, written in Latin script; it represents a good example of the old stage of the southern Tosk dialect; the book is very valuable for the study of both the Italo-Albanian dialect and Arvanitica; see Sasse 1991: 8–9) and flourished in the 19th century.

The period starting from the middle of the 19th century marks the beginning of the development of the Standard Albanian language.

*The ethnic and language name.* The old Albanian ethnic and language name has been used since the oldest Byzantine and Latin sources in the two forms of the root *arb-* and *alb-* (which form is etymologically primary remains unclear): *Ἀρβανιτῶν* by Attaleiates, *lingua Albanesca* in the Dubrovnik document from the 13th century and so on (cf. also *Ἀλβανοί* and *Ἀλβανόπολις* by Ptolemy; see below). Albanians themselves used this name only with the root *arb-* (see, however, the name of Albanian tribe *lab*, and of the corresponding area *Labëri*, probably from *alb-* with metathesis). The name with the root *arb-* is mentioned in old Albanian documents, but it went out of use in the main part of Albanian-speaking

area and remains in use only in diaspora dialects (It.-Alb. *arbëresh*, Gr.-Alb. *arvanitas*). In other areas, it has been replaced by the term with the root *shqip*–: *shqip* ‘Albanian (used as an adverb)’, *gjuha shqipe* ‘Albanian language’, *shqiptar* ‘Albanian (ethnic name)’ and so on. This name is derived from the verb *shqiptoj* (‘to speak clearly’, earlier ‘to understand’; see B. Demiraj 2010); it was first registered in the 16th century (at the beginning in its adverbial form) and consolidated in the 18th century. The ethnic name change was triggered most likely by the Ottoman conquest (end of the 14th century–second half of the 15th century), which led to deep changes in the structure of the Albanian nation (for the Albanian ethnic and language names see, first of all, B. Demiraj 2010).

*The origin and early history of the Albanian language.* The fact that written evidence concerning the Albanian language and Albanians is historically rather shallow makes it problematic to reconstruct the earlier stages of Albanian language history. From the end of the 19th century, a central place in this problem belongs to the question of the possible development of Albanian from one of the languages that were spoken in Antiquity in the northern part of the Balkans. However, this question itself resulted from speculations that were largely deductive. Two main theories consider Albanian as a descendant of either Illyrian or Thracian languages, respectively (for an overview see Matzinger 2009, 2012b). According to the former theory, Albanians were autochthonous to the area that roughly coincided with their modern territory; the latter theory assumes that the ancestors of Albanians came to the modern territory from a more eastern part of the peninsula. This aspect of the “origin problem” accounts for the political shade of the debate: the Illyrian theory became a sort of a sacred cow in Communist Albania and remains the mainstream theory among Albanian scholars nowadays. The situation is complicated by the fact that the exact extent of the idioms referred to as the Illyrian and Thracian languages, respectively, is not known. In the case of Illyrian, the issue is relevant, first of all, for Messapian; the status of this idiom is indeterminate and varies in linguistic writings from merely an Illyrian dialect to a different language. In the case of Thracian, one should mention “Daco-Misian,” which was assumed by V. Georgiev to be a separate language and a direct ancestor of Albanian. In the last decades, Dardanian is considered a separate language, also potentially important for Albanian language history.

There are some arguments in favour of either theory. From a purely *historical* point of view, the absence of the mentions of Albanians till the 11th century cannot be considered as an argument against (or in favour of) either the “autochthony” or “migration” hypotheses. The mentions of the tribe *Ἀλβανοί* and the city (?) *Ἀλβανόπολις* (to the East of the Dyrrachion) in Ptolemy’s *Geography* (2nd century A.D.; see also *Ἄρβων*, a city in Illyria with uncertain localisation by Polybius, 2nd century BC) don’t testify definitively in favour of the historical continuity of the Albanians in this territory: there are many known examples in world history where an ethnic or a language name shifted from one ethnos to another.

There are studies that show some kind of *archeological* continuity between Antiquity and the second half of the first millennium A.D., first of all the so-called *Koman culture* near Kruja. However, regardless of the particular interpretation of excavation materials from an archeological point of view, these studies cannot be decisive for the solution of the problem of ethnic and linguistic continuity.

There are also some *linguistic* data that can be taken as evidence in favour of the Albanian language’s descent from either the Illyrian or Thracian language, respectively. Unfortunately, most of these arguments are weakened by the insufficiency of our knowledge about the ancient Indo-European languages of the Balkan Peninsula.

There are some reliable etymological correspondences between Albanian words and both Illyrian (including Messapian) and Thracian language material (a few glosses and



numerous onomastic tokens). Among the most commonly accepted Illyrian-Albanian correspondences are Illyr. ῥινός ‘cloud, fog’ – Alb. *re, rë* ‘cloud’; Messap. βρένδον ‘deer’, βρέντιον ‘deer’s head’ – Alb. *bri, brī* ‘horn’; Messap. *Juppiter Menzana* ‘the God to whom horses were sacrificed’ – Alb. *mëz, mǎz* ‘foal’; Tarentian βάρυκα:αἰδοῖον – Alb. *bark* ‘belly’; Messap. *aran* ‘field’ – Alb. *arë* ‘field’; Messap. *bilia* ‘daughter’ – Alb. *bijë* (dial. *bilë*). The Thracian-Albanian (and Dacian-Albanian) correspondences include Dac. μαντεία ‘blackberry’ – Alb. *man* ‘mulberry’; \*δράντις (Hesychius) ‘deer’ – Alb. *drë* ‘deer’ and some other.

Lexical data of this kind are obviously insufficient for any kind of decisive conclusions about the relations between Albanian and the two ancient Balkan languages. The same is also true of phonetic Illyrian-Albanian and Thracian-Albanian isoglosses. Given the limits of our knowledge, it is preferable to consider Illyrian, Thracian and Albanian as separate branches within the Indo-European language family.

There are, however, other pieces of linguistic evidence that may be relevant for the determination of the place of the “intermediate Balkan Urheimat” of the Albanians, regardless of whether or not the Albanian language originates from the Illyrian language.

- One piece of evidence comes from the phonetic form of some ancient geographic names localised in the modern Albanian territory, first of all on the Adriatic coast. If it is possible to trace a linguistically continuous path in the phonetic development of these names from their ancient form to their modern form in accordance with Albanian “sound laws,” it would support the hypothesis that the Albanians have been living on their modern territory from Antiquity. There are, however, some phonetic difficulties in reconstructing such a continuous development, for example, with respect to such important Albanian toponyms as *Shkodër* < *Scodra*, Σκόδρα, *Durrës* < *Dyrrhachium*, Δυρράχτιον, *Shkumbi(n)* < *Scampis*, *Scampinus*. There is a long discussion on the historical development of these names; the most widespread view has it that the historical development of these names does not coincide with the laws of the Albanian historical phonology (e.g. *Scodra* should have developed into something like \*Hadër; see Matzinger 2009: 23). By contrast, in the inner Balkans there are some geographic names that show a phonetic development according to Albanian “sound laws” (see e.g. *Nish* < *Naissus*, Ναισσοῦς).
- Albanian has a rather modest number of old Greek borrowings (33 according to Ölberg 1972). If the ancestors of Albanians lived near the Adriatic coast and communicated with the inhabitants of Greek colonies (*Dyrrhachium*, *Apollonia* et al.), the number of Greek loans would have been higher.
- Albanian lacks almost any inherited lexemes related to seamanship. We know, though, that the Illyrians, who were inhabitants of the seacoast, were artful seamen and pirates.

All these pieces of evidence indicate, although somewhat indirectly, that in the époque of late Antiquity the ancestors of Albanians did not live on the Adriatic coast or in close proximity to it. We can say nothing, however, about where they actually did live at this time.

Another group of linguistic facts is more relevant for the solution of the Albanian homeland problem. These are the numerous Albanian-Romanian correspondences. These correspondences pertain to various linguistic levels. There are numerous correspondences in grammar, first of all a striking similarity in the architecture of the noun phrase (see the sections “Morphology” and “Syntax”). The source of the common development is, most likely, Albanian. There are also some phonological (e.g. diphthongization of *e*, see

the section “Phonology”) and morphological (similar patterns of number expression in masculine nouns; see the section “Morphology”) resemblances. Numerous Albanian-Romanian lexical correspondences include early Albanian borrowings into Romanian, those words that entered both languages from some common (and unknown) sources as well as Latin words that are found exclusively in Romanian and Albanian (see Vătăşescu 1997, Kaluzhskaya 2001). These phenomena point to a long period of very intensive contacts between the ancestors of Albanian and Romanian. The time of these contacts should coincide with the period of the formation of the (Daco)Romanian language, that is, the middle of the first millennium A.D. As far as localization is concerned, we may hypothesize that it was the inner part of the Balkan Peninsula (maybe the territory of ancient Dardania; see Jokl 1924; the triangle Niš–Sofia–Skopje, see Weigand 1927).

*The layers of the Albanian lexicon.* It is believed that Albanian belongs to languages with a high amount of lexical borrowings (in this respect Albanian is similar to Armenian; see Matzinger *in print*). The etymological dictionary of inherited (i.e. non-borrowed) Albanian words by Bardhyl Demiraj lists 572 tokens (without the derivatives that originated in Albanian; B. Demiraj 1997: 37). Since Albanian is a language with a relatively shallow written attestation, the importance of borrowings as data for the study of historical phonetics (and, to a lesser degree, historical morphology) is extremely high.

In terms of chronology, the earliest massive layer of borrowings in Albanian is the *Latin* layer (for a small number of old Greek loans, see above). The period of intensive Albanian-Latin language contacts might have begun in the early 1st century AD, after the final incorporation of the West Balkans into the Roman state (Romans’ first contacts with the West Balkan seacoast happened in the 3rd century BC) and continued until the 5th–6th century AD, eventually taking the form of Albanian-(Proto)Romanian contacts. The Latin borrowings in Albanian (see Mihăescu 1966, Haarmann 1972, Landi 1989, Vătăşescu 1997, Bonnet 1998) are very numerous: according to the various lists the number of Latin etymons in Albanian is no less than 600; these borrowings cover practically all parts of the Albanian lexicon in both semantic and morphological dimensions. It should be noted that although the main part of Latin loans belongs to the cultural lexicon (see e.g. *gjyq* ‘trial, court’ < *iudicum*; *mjek* ‘physician’ < *medicus*), there are also some basic vocabulary items that have been borrowed from Latin (see e.g. *vij* ‘come’ < *venio*; *këmbë* ‘leg, foot’ < *\*camba* et al.). There are some derivational affixes borrowed from Latin (see the section “Word Formation”). Any direct influence of Latin on Albanian phonetics and morphology is rather uncertain. The Latin borrowings came into (Proto)Albanian at the time when (Proto)Albanian was experiencing great changes, and that is why these borrowings constitute the main source of our knowledge about the absolute chronology of Albanian language history.

The Albanian-Slavic language contacts (see Selishchev 1931, Jokl 1934, Desnickaja 1968, Svane 1992, Ylli 1997) began after the Slavic invasion into the Balkans and continue, in a sense, until now. Unlike Latin borrowings, which are for the most part common Albanian, a great part of Slavic loans have a clear dialect distribution (according to Ylli 1997 only a quarter of some 1000 Slavic borrowings have a more or less pan-Albanian distribution). According to historical-phonetic criteria, there is a small group (circa twenty) of Slavic borrowings that had to enter Albanian relatively early (before the 10th–11th century), whereas the chronological fixation of the most part of Slavic loans remains unclear. The main source of the Slavic loans for the southern dialects of Albanian was Macedonian-Bulgarian dialects, while the main source for the northern dialects was Serbo-Croatian dialects. Slavic borrowings belong to various semantic domains, but are mostly related to the cultural lexicon (*oborr* ‘yard’, cf. Bulg.-Mac. *obor*, SCr. *obor*; *bujar* ‘noble’ < Slav. *boljar*). Some Slavic derivational suffixes also entered Albanian.

The Middle and Modern *Greek* borrowings (prevailing in southern Greek dialects of Albanian) date back to early medieval times. *Italian* loans date back to the beginning of active contacts by the Italian states with this part of the Dalmatian coast (11th century); many early Italian borrowings came from the Venetian dialect.

After the Ottoman invasion (end of the 14th–15th century) a great influx of *Turkish* words (including Arabic and Persian borrowings in Turkish) overflowed Albanian. In this layer, there are a great number of terms related to economic, administrative and spiritual life (*bajrak* ‘banner, an administrative district in the Albanian mountains’ < *bayrak*, *borxh* ‘debt’ < *borç*, *sevda* ‘love’ < *sevda*), but also various pragmatically loaded, “gesture-like” words (*kismet* ‘fate (used as an interjection)’ < *kismet*), as well as some words that belong to basic vocabulary (*jeshil* ‘green’ < *yeşil*). Since the beginning of the Albanian national-cultural revival (the so-called National Renaissance or National Awakening) in the second half of 19th century the main trend in Albanian language policy was to eliminate Turkish loans as a symbol of national oppression and undesirable “easternization”. Many Turkisms really went out of use, and some shifted to the domain of popular, “sub-standard” language, but many Turkish words remain part of the present Albanian lexicon. Albanian also has a small amount of grammatical morphemes borrowed from Turkish (on Turkish borrowings see, first of all, Boretzky 1975–1976).

Since the end of the 19th century and, especially, in the 20th and 21st centuries, an innumerable amount of *internationalisms* entered Albanian. The sources of these words depended on political circumstances (Italian and French before the Second World War, Russian in the first two decades of Communist rule, and English in recent decades).

Albanian diaspora dialects have a great number of words from current and old neighboring languages: Arvanitica from Greek, Arbresh from Greek and Italian (a big part of the ancestors of present Italo-Albanians came from Greece) and so on.

*Areal position of Albanian.* There have been numerous attempts to determine the areal place of Albanian within the IE languages (see, for example, Porzig 1974, Pisani 1950, 1959, Jokl 1963, Hamp 1966, Orel 2000). The detailed investigation of lexical as well as, to a lesser extent, phonetic and grammatical isoglosses indicates that there are especially close connections of Albanian with North-Eastern IE languages (Baltic and Slavic), on the one hand, and with Greek, on the other hand. There are discrepant opinions with respect to the chronological assessment of these two groups of convergences.

Recently, there has been a discussion of the so-called Balkan Indo-European, which includes Greek, Albanian, Armenian and Phrygian, and which is understood as an intermediate protolanguage (see the overview in Matzinger 2012a).

Modern Albanian is considered one of the base members of the Balkan Sprachbund, which also includes Bulgaro-Macedonian, some Serbian dialects, Balkan Romance, modern Greek and Romani. Standard Albanian indeed shares the majority of the so-called Balkanisms, such as the presence of stressed /ə/, the postpositive article, the merger of genitive and dative, the lack of the infinitive, the volitive future, the object clitic doubling and some others (see “Phonetics”, “Morphology” and “Syntax”; see also Schaller 1975, Asenova 2002). The Balkanisms must have been formed in the broad temporal interval between the middle of the first millennium A.D. and the first centuries of the Ottoman invasion, which covers the second half of the Proto-Albanian and the Old Albanian period. The same time span is the period of the europeanization of Albanian: Albanian shares 8 or 9 of 12 diagnostic features of Standard Average European as listed in Haspelmath 2001 (among them the presence of both definite and indefinite articles, the possessive perfect, relative clauses with relative pronouns, the participial passive etc.). Apparently, balkanization and europeanization of Albanian could have developed parallel to each other.

*The Albanian alphabet.* The modern Albanian alphabet, which was elaborated and accepted in 1908, is consistently phonological. It consists of 36 graphemes (including digraphs). The following graphemes should be explained:

**TABLE 12.1 ALBANIAN ALPHABET**

grapheme	phonemic value
c	ts
ç	č
dh	ð
ë	ə (stressed and unstressed, sometimes it isn't pronounced)
gj	J
l	l'
ll	l
nj	n'
q	c
rr	r:
sh	š
th	θ
x	dz
xh	ž

### **Periodization of Albanian language history, reconstructed stages of Albanian**

In the following sections I shall use the following nomenclature for periods in Albanian language history.

- *Pre-Proto-Albanian period*, the language of the ancestors of Albanians before the beginning of intensive contacts with Rome, i.e. before the first century AD. In the end of this period many important sound changes took place that moved (Proto)Albanian away from the (late) PIE state, e.g. the merger of PIE \*o and \*a as well as the merger of PIE \*ā and \*ē, considerable changes of the consonant system (the merger of the PIE Media and Media Aspirata, the changes in the gutturals, which, however, have preserved an opposition of the reflexes of all three series); and, probably, the stress changes (see the section “Phonology”).
- *Proto-Albanian*, the language of the first millennium AD, the time of intensive contacts with Latin and early contacts with Slavic. A new wave of important changes took place during this period, among them the reduction of the pre- and post-tonic syllables (which drastically changed the phonetic shape of Albanian word); the old quantitative vowel opposition was lost; the intervocalic voiced consonants were weakened and, partly, lost (the exact conditions of this loss are not completely clear; see the section “Phonology”); as a result of the syllable reduction a large part of PIE endings were lost, and eventually many new morphological formants were agglutinated instead; important grammatical balkanisms arose in this period (most likely due to intensive contacts between the ancestors of the Albanians and Romanians; see the sections “Morphology” and “Syntax”). At the end of this period, just before the split into the Tosk and Gheg dialects, that is, at the end of the first millennium AD, the (Proto) Albanian language took a form that is quite similar to that of the Albanian language as known to us from written sources.

- *Old Albanian*, the language from the time of the first dialect split till the end of the period of the old written texts, that is, till the 18th century. This period can be further subdivided into Early Old Albanian (from the beginning of the first millennium AD to the earliest written texts) and Late Old Albanian (16th–18th centuries). Early Old Albanian was characterized by secondary contacts between the Gheg and Tosk dialects in the territories that roughly coincide with the modern Albanian-speaking area and by the processes of forced balkanization of Albanian, especially in the south of the Albanian-speaking territory. In the Late Old Albanian period (the time after the Ottoman conquest), the modern Albanian dialectal landscape was formed.
- *Modern Albanian*, the Albanian language of the 19th–21st centuries.

This periodization presented here is almost identical to those used by B. Demiraj (1997), Matzinger (2006), Hock (2005, with an analytical overview of previous attempts at periodization), Schumacher/Matzinger (2013) and de Vaan (*in print*) and differs somewhat, mainly in terminological details, from those by Desnickaja (1984), Ölberg (2013) and Orel (2000).

In the sections considering phonetics and grammar I shall operate with reconstructed Albanian units (phonemes, fragments of the phonemes systems, forms, paradigms and so on). The following designations will be used:

- *PA* for the stage that is reconstructed based on the comparison of the Gheg and Tosk dialects as well as the language of the Old Albanian written texts. It can be assumed that this reconstructed stage corresponds more or less to the Albanian language just before the dialect split, i.e. to Albanian at the very end of the Proto-Albanian period.<sup>1</sup> Sometimes the designation EPA, early Proto-Albanian, is used for units that belong to relatively earlier stages in the development of Albanian within the Proto-Albanian period.
- *PPA* (only when needed) for the stage of the Albanian development just before the beginning of the contacts with Rome, i.e. on the border between the Pre-Proto-Albanian and Proto-Albanian periods.

In addition, the term Standard Albanian (SA) will be used.

## PHONOLOGY

In this section I chiefly rely on the following surveys of Albanian historical phonology: B. Demiraj 1997, Matzinger 2006, Ölberg 2013 (for vocalism); Schumacher/Matzinger 2013; and de Vaan *in print*; I also take into account Çabej 1976; Sh. Demiraj 1996, Orel 2000, Topalli 2006; see also the overviews of the development of Albanian sound system in Hock 2005 (vocalism and consonantism) and Vermeer 2008 (vocalism).

In what follows the vowel, resonant and consonant subsystems of Albanian as well as the problem of Albanian stress patterns will be described in turn. For each subsystem, a discussion of the SA system will be followed by the reconstruction of the alleged PA stage, based on the data from dialects and the earliest written texts. The history of PIE sounds in Albanian will be the topic of the next part in each paragraph. Only the main lines of the sound development will be specified; for the others the necessary references will be given. The illustrative etymological material (based, first of all, on B. Demiraj 1997, Schumacher & Matzinger 2013, Orel 1998 and Çabej 1976–2014) will be given in a very selective way.

## Vowels

### SA vowel system

i	y			u
	e	ə	o	
		a		

All Albanian vowels may be used in both stressed and unstressed positions. Therefore, the standard Albanian vowel system is very close to the idealized Balkan vowel system (five basic vowels + stressed ə): the only extra vowel is the labialized high front phoneme *y*.

### *Vowels: basic dialectal differences. Reconstruction of the PA system*

The Standard Albanian system above is very close to the system attested in the majority of Tosk dialects. Other dialects differ, however, in this respect from the standard language to a great extent. The major differences are the following.

- In the Gheg dialects and in a small part of the Tosk dialects (namely in their southern parts, that is, the Lab and Cham dialects, as well as in some Arbresh dialects) there is an opposition between long and short vowels. This opposition is also attested in old Albanian written texts; thus, it used to be a pan-Albanian feature that was later lost in the greater part of the Tosk area.
- The Gheg dialects have a special series of nasal vowels (*ã, ï, ù, ÿ, ë*, but nasal *õ* is lacking), which developed in the PA period and disappeared in the Tosk dialects. It should be noted that the authors of the first written texts either do not mark vowel nasalization at all or do so in a very inconsistent way. PA *ã* developed into Tosk *ë /ə/*: Gheg *ãsht* '(s)he is' – SA, Tosk. *është* id.; PA *ë* developed into *ë* or *e* depending on the dialect (and depending on the phonetic position in some dialects) in Tosk: Gheg *vënd* 'place' – SA, Tosk *vend*, *vënd* id. Whether there is/was a phonological opposition between *long nasal vowels* and *short nasal vowels* in Albanian dialects and in PA remains an unsolved problem.
- SA lacks diphthongs as independent phonological units; all vowel sequences within words are treated as biphonemic combinations. During the PA period the processes of diphthongization were at work in Albanian, which led to the arising of three diphthongs, namely *uo*, *ie* and *ye* (on the origin of these diphthongs see below, section 1.3). These diphthongs later had different patterns of development in Albanian dialects:

*uo* > *ua* in Tosk: *gr'ua*, *gru'a* 'wife' (with different stress patterns depending on the phonetic context and dialect); *ue*, *ū* more rarely *ua* and even *uo* in various Gheg dialects: *grue*, *grū* 'id.';

*ie* > *ie* in Tosk (with different stress patterns depending on the phonetic context and dialect): Tosk *d'iell*, *dj'ell* 'the Sun'; *i*, more rarely *ie*, in Gheg dialects: Gheg *dill* 'id.';

*ye* > *ye* (with different stress patterns depending on the phonetic context and dialect), sometimes *ie* in Tosk: *kr'ye*, *kry'e*, *krie* 'head'; *ye*, *ÿ* in Gheg dialects: *krye*, *krÿ* 'id.'

- The number of vowel phonemes differs in various dialects. The major differences (along with the presence/absence of long/short and oral/nasal oppositions) are (see Gjiniari et al. 2007):

- A) absence of *y* (> *i*) (as well as its correlates  $\bar{y}$  and  $\bar{y}$ ) in some Albanian dialects, both Tosk (Lab and Cham dialects) and Gheg;
- B) In Gheg dialects only unstressed  $\bar{e}$  / $\bar{\partial}$ / is present. During the last centuries a process of the reduction of unstressed  $\bar{e}$  has been under way in Albanian. This process is especially strong in the Gheg dialects. It seems that unstressed  $\bar{e}$  is not phonemic in the majority of Gheg dialects. By contrast, analysis of the old Gheg texts points toward the phonemic status of unstressed  $\bar{e}$ . In Lab and Cham dialects there is an opposition between / $\bar{\partial}$ / and / $\bar{\partial}$ :/.
- C) In northern Gheg dialects, there are two additional long vowel phonemes with very restricted spheres of use (only some words for each).
- / $\bar{\alpha}$ /, / $\bar{\epsilon}$ / < *ae* (*ae* is used as such in the earliest written texts); this vowel is typical in a large part of the northern Gheg and central Gheg dialects: *th/* $\bar{\alpha}$ / ‘you said (aor.)’, other dialects: *the* ‘id.’
  - / $\bar{\delta}$ / < *oe* (*oe* is used as such in first written texts); this vowel is found in some western northern Gheg dialects: *v/* $\bar{\delta}$ /*s* $\bar{e}$  ‘dew’, other dialects: *ves* $\bar{e}$  ‘id.’
- In the central Gheg dialects, a set of diphthongization processes (as well as some other vocalic changes) took place, probably in the Late Old Albanian period: /*i*/ > /*aj*/, /*ej*/, /*aj*/, /*e*/; /*u*/ > /*ou*/, /*au*/; /*a*/ > /*æ*/; /*o*/ > /*œ*/, /*ou*/; /*y*/ > /*i*/, /*aj*/, /*œ*/; the long and nasal correlates of these vowels followed parallel patterns of development. For the phonetic conditions of such changes, see Jusufi 2011; the phonemic status of the variants that emerged in this process is not entirely clear.

To sum up, the following vowel system can be reconstructed for the *very end of the PA period*:

i ī i y ŷ ȳ	e ē ē (ə) o ō	u ū ũ i ī i	ie ye uo
	a ā ã		

Notes: / $\bar{\partial}$ / appeared only as an unstressed vowel.

The opposition of the short and long vowels might have arisen either at the end of the PA or the very beginning of the OA period. In the latter case the opposition must have developed in the Gheg and Tosk zones independently under similar conditions, which is less plausible. It remains unclear whether there was ever an opposition between short and long vowels in unstressed syllables.

Perhaps there also existed a special series of long nasalized vowels.

The system as reconstructed above is very close to that of some conservative western northern Gheg dialects.

### 1.2 Vowels: the origin

The origin of PA long and nasalized vowels will not be traced in this section (see section 1.3 on their origin).

TABLE 12.2 PIE SHORT VOWELS (IN STRESSED POSITION)

PIE phonemes	PA and SA reflexes	examples
1. *a (also *h <sub>2</sub> e), *o (also *h <sub>3</sub> e), *H (in the roots in zero-grade)	PPA, PA *a > SA <i>a</i>	SA <i>bathë</i> ‘broad bean’ < PIE *b <sup>h</sup> ak-; SA <i>natë</i> ‘night’ < PIE *nok <sup>w</sup> t-; SA <i>mas, mat</i> ‘to measure’ < PIE *mHt-(io)-
2. *e (also *h <sub>1</sub> e)	a) PA, SA <i>e</i> after CR- and/or before-NC; b) PA, SA <i>ja</i> ; c) PA, SA <i>iel_*</i> r, <i>l</i> d) PA, SA <i>je</i> (elsewhere)	a) SA <i>breth</i> ‘to wander’ < PIE *b <sup>h</sup> red <sup>(h)</sup> -; <i>dhemb</i> ‘to ache’ < PIE *ǵemb <sup>h</sup> - b) SA <i>jam</i> ‘I am’ < PIE *h <sub>1</sub> esmi c) SA <i>bie</i> ‘to bring’ < PIE *b <sup>h</sup> er- d) SA <i>djeg</i> ‘to burn’ < PIE *d <sup>h</sup> eg <sup>wh</sup> -
3. *i (also *Hi)	PA, SA <i>i</i>	SA ( <i>i</i> ) <i>lig</i> ‘bad, ill’ < PIE *h <sub>3</sub> lig-
4. *u (also *Hu)	PA, SA <i>u</i>	SA <i>gjumë</i> ‘sleep’ < PIE *sup-no-

## Notes

- Latin *o* in borrowings manifests as *o* in SA; Latin *a, ā* > Alb. *a*. Based on this evidence, the change from PIE \*o to \*a can be dated back to the PPA period.
- Latin “wide” *e* (< Lat. *ĕ, æ*) experienced similar changes in roughly identical phonetic conditions also, so we may attribute them rather to the end of the PA period.
  - The conditions are not entirely clear.
  - Maybe the change has a later character and is parallel with the diphthongization of other PA vowels (see 1.3).
- 4. Lat. *i* > SA Alb. *i* (on the reflex Lat. *i* > Alb. *e* see below); Lat. *u* > SA *u*.

## PIE long vowels

The vowels in the PIE lengthened grade, the sequences \*Vh<sub>1/2/3</sub> and vowels resulting from the sequences \*Vsr/l (with the following loss of the \*s: SA *dorë* < PIE \*ǵ<sup>h</sup>es-r-) must have merged in the PPA period.

TABLE 12.3 PIE LONG VOWELS (IN STRESSED POSITION)

PIE phonemes	PA and SA reflexes	examples
1. *ā (and *eh <sub>2</sub> ), *ē (and *eh <sub>1</sub> )	PPA *ā > PA *o (earlier *ɔ) > SA <i>o</i>	SA <i>motër</i> ‘sister’ < PIE *meh <sub>2</sub> -tr-; SA <i>mot</i> ‘time, weather’ < PIE *meh <sub>1</sub> -t
2. *ō (and *eh <sub>3</sub> , *oH)	PPA *ō > PA <i>e</i> (earlier *æ), SA <i>e</i>	SA <i>derë</i> ‘door’ < PIE *d <sup>h</sup> wōr-
3. *ī (and *iH)	PA, SA *i	SA <i>pi</i> ‘drink’ < PIE *pih <sub>3</sub> -
4. *ū (and *uH)	PA *y > SA *y; later *y > <i>i/_(s)#</i>	SA <i>gjysh</i> ‘grandfather’ < PIE *suH-so-; SA <i>mi</i> , Ghëg. <i>mī</i> ‘mouse’ < PIE *mūs

## Notes

- Old Greek borrowings in Albanian show the replacement of *ā* by SA *o*. Lat. *ā* gives Alb. *a* (for another opinion, see, however, Schumacher & Matzinger 2013: 221), and Lat. *ē* > Alb. *e, i*. Based on this evidence, the merger of PIE vowels and their change to a vowel other than *a* must have taken place at the end of the PPA period.
- Latin *ō* also changed to *e*, at least in the older group of Latin borrowings, which makes it possible to establish a *terminus ante quem* for this process. The intermediate stage of the development of this vowel should have been *æ*. Another reflex of Latin stressed *ō* in Albanian is *o*; most likely, this reflex is found in the later layer of borrowings – it coincides with the reflex of the Latin stressed short *o*. Latin short *o* in Albanian also has the reflex *u* (maybe in the older borrowings).



It can be hypothesized (see Matzinger 2012b) that at the beginning of the contacts with Latin, that is, at the beginning of the PA period, the early PA vowel system still had quantitative distinctions. Thus, the Latin long  $\bar{o}$  was replaced by early EPA  $*\bar{o}$  (later  $*\bar{e} > SA e$ ), Lat. short  $o$  by PA  $*u$  (short  $*o$  was lacking in the early PA system after the change from PIE  $*o$  to PPA  $*a$ ), and both Lat.  $a$  and  $\bar{a}$  by PA  $*a$ . Then, after the change  $*\bar{o} > *\bar{e}$ , both Lat.  $o$  and  $\bar{o}$  (which have lost their quantitative distinctions in Late Latin and merged in East Romania) came to be replaced by PA  $*\bar{u}$  (< PIE  $*\bar{a}$ ,  $*\bar{e}$ ).

3. Lat.  $\bar{i}$  reflects as SA  $*i$  (SA *mik* ‘friend’ < Lat. *amicus* ‘friend’). Latin short  $i$  had two reflexes: SA  $i$  and SA  $e$ . The same two reflexes are also characteristic of Lat.  $\bar{e}$ . It seems that at the time of the contacts with Albanian, Lat.  $i$  and  $\bar{e}$  merged in the close  $*\bar{e}$  (it was characteristic of Vulgar Latin in general), which differed from open  $*e$  (<  $e$ ,  $ae$ ). In the first period of the contacts (when PIE  $*ay$ ,  $*oy$  hadn’t yet changed into  $*\bar{e} > SA e$ ; see below, and also de Vaan 2004) Lat.  $*\bar{e}$  was replaced by PA  $i$ ; later, when this change was already in progress, Lat.  $e$  began to be reflected by  $*\bar{e}$ .
4. Lat.  $\bar{u}$  has two reflexes in Albanian: SA  $y$  (older layer) and SA  $u$ .

### PIE diphthongs

There is almost no reliable etymological data concerning the development of PIE long diphthongs in Albanian.

**TABLE 12.4 PIE DIPHTHONGS (IN STRESSED POSITION)**

PIE phonemes	PA and SA reflexes	examples
1. $*ay$ , $*oy$	EPA $*\bar{e} > PA *e > SA e$	SA <i>edh</i> ‘kid’ < PIE $*(H)ay\acute{g}$ -; SA <i>shteg</i> ‘path, road’ < PIE $*stoyg^h$ -
2. $*ey$	PA, SA $i$	SA <i>dimër</i> , Gheg <i>dimën</i> ‘winter’ < PIE $*\acute{g}^h\bar{e}ym\bar{o}n$
3. $*aw$ , $*ow$	PA, SA $a$	SA <i>ag</i> ‘dawn’ < $*h_2ewg$ -; SA <i>laj</i> ‘to wash’ < PIE $*lh_3\text{-}ew$ -
4. $*ew$	EPA $*\bar{e} > PA *e > SA e$	SA <i>hedh</i> ‘to throw’ < PIE $*skewd$ -

#### Notes

3. Lat.  $au > SA a$  in earlier borrowings, and  $> af$  in later borrowings.

### “Inner Albanian” development of the vowels in the PA period

#### a) “Umlaut” ( $i$ -mutation)

EPA  $*a$  (<PIE  $*a$ ,  $*o$ ,  $*H$ ) > PA, SA  $e$ , if the following syllable has  $i$  (or  $\bar{i}$ ). This change affected inherited words and the words borrowed from Latin, but not the words of Slavic origin (a few exceptions are usually explained away as a result of analogical processes). This process had a great significance for the Albanian morphological system (see the section “Morphology”): SA *net* ‘nights’ < EPA  $*na(k)ti(h)$ . The “umlaut”  $e$  does not have diphthong-like reflexes. The umlaut pattern EPA  $*o > PA, SA e$  and EPA  $*u > PA, SA y$  occurs in the same conditions but is less widespread.

The conditions of the change PA  $*e > i$  in some verbal and nominal forms are not clear ( $i$ -mutation, Orel 2000: 145, de Vaan 2004: 71; change before the following consonant cluster, Hamp 1971: 224; the influence of the following palatal consonant, Schumacher & Matzinger 2013: 219–220). The significant role that the alternation  $e : i$  has assumed in the OA and SA verb system (see below, the section “Morphology”) could have been propelled by analogical processes.

## b) Diphthongization

EPA \*o (< \*ɔ, also < Lat. *o*, *ō*) > \*uo/ \_\*r#, \*l#, \*n#; \*n in this position disappeared: PA \*kron (< Gr. κρήνη, Dor. κράνα) > SA *krua* (Gheg *krue*, *krū*). In the anlaut \*o (< \*ɔ) yielded Gheg *vo-*, Tosk, SA *va-* (SA *vatër*, Gheg *votër* ‘hearth, fireplace’ < \*h<sub>2</sub>eh<sub>2</sub>-tr-).

EPA \*œ > *ye/* \_\*r#, \*n#; \*n# disappears obligatorily, \*r disappears optionally (see Schumacher & Matzinger 2013: 224, SA *arsye* ‘cause’ < Lat. *ratiōnem*).

On the development EPA \*e > PA, SA *ie/* \_\*r#, \*l# see above.

## c) Development of the unstressed vowels

One of the most important phonetic changes in the PA period was the reduction of unstressed vowels. Two remarks that concern the chronology of this process should be made: (a) this process affected most Latin loans (see *gusht* ‘August’ < Lat. *augustus*); and (b) it continued to operate after umlaut processes happened (SA *dash* ‘ram’ < \*deš-i < \*-oi). These facts indicate that the reduction of unstressed vowels was at least partially under way at the end of the PA period. These processes yielded a very complicated and hardly reconstructable picture. In the absolute anlaut unstressed vowels were lost (see *gusht* above), in word-internal positions (both pre-tonic and post-tonic) they were lost in some words (SA *mbesë* ‘niece’ < Lat. *nepōtia*) and got reduced to /ə/ in other words (SA *këndoj* ‘to sing, to read’ < Lat. *cantāre*), and it is impossible to establish the strict rules that governed this distribution. As for the auslaut vowels, it seems that the process of reduction was influenced by morphological conditions (see the section “Morphology”). Generally speaking, it could be supposed that the reduction of unstressed vowels was a rather long-term process: it presumably began as early as in the Pre-Proto-Albanian period and continued at least until the end of the Old Albanian period. A more detailed analysis with examples can be found in de Vaan *in print*; B. Demiraj 1997: 46–47.

## d) Emergence of nasalization

In the last phase of the PA period a series of nasal vowels developed out of VN sequences with possible loss of \*-n in the auslaut syllables (see also 1.1). The vowel *o* did not undergo nasalization and changed instead into the diphthong *ua* before nasal consonants.

## e) Emergence of the new quantitative opposition

The new quantitative opposition arose, it seems, at the end of the PA or the very beginning of the OA period. The main sources of the lengthening were various processes of contraction, which involved both inherited IE words and Latin borrowings in Albanian (see e.g. *bē* ‘oath’ < \*b<sup>h</sup>eyd<sup>h</sup>ā; *mjek* ‘physician’ < Lat. *medicus*). Lengthening was also caused by the influence of the following resonants *-r* and *-l*. Although dialectally restricted, the processes of lengthening intensively continued, however, later in the Old Albanian period; e.g. vowel lengthening could occur as a result of apocope: eastern-northern Tosk *plak/əl* ‘old woman’ – Gheg */pla:k/* ‘id.’

*The development of the vowel system. Conclusion*

To sum up, the main trend of the development from the PIE vowel system into the Albanian vowel system was the loss of the IE quantitative distinctions and their partial replacement

by qualitative distinctions. This process took a long time. Some of the important changes that disrupted the symmetry of the long and short vowel subsystems had occurred already in the PPA period (the merger of PIE \*a and \*o, the merger of the PIE \*ā and \*ē). At the beginning of the PA period, the “old” quantitative opposition was still present in Albanian. Very important vocalic changes took place in the PA period. Reflexes of Latin borrowings make it possible to establish a relative chronology of these changes (see e.g. de Vaan 2004). Among the most important PA changes are the development of \*ū into \*y, the change \*ā > \*ɔ > \*o, and the development of various reflexes of PIE \*e. At a certain moment the PA system possessed two *e*-like phonemes: the open \*e (< PIE \*e) and close \*ē (< PIE \*ay, \*oy, \*ew). The second one did not have diphthong-like reflexes. The presence of these two phonemes was paralleled by the Balkan Latin vocal system, which also had two *e*-phonemes (and only one /o/ phoneme, which is also true in Albanian). Whether there was a direct Latin (Roman) influence on Albanian phonetics in this respect remains, however, unclear. Later in the PA period the processes of the diphthongization before sonorant phonemes took place (\*e > *ie*; \*o (or \*ɔ) > *uo*; \*æ (< PIE \*ō) > *ye*) as well as unlaut processes. The vowel *e*, resulting from the *i*-fronting of *a* (as well as from \*æ < PIE \*ō), merged, at the end of the PA period, with the identical reflexes of \*e and \*ē. Thus, a new quantitative opposition emerged; finally, nasalized vocal allophones gave rise to a special series of phonemes.

### Sonorants

#### SA sonorant system

m	n	n'
	l	l̥
	r	r:
	j	

#### *Sonorants: main dialectal differences. Reconstruction of the PA system*

- The absence of the opposition *r* : *r*: (orthographic *rr*) in some dialects (both Gheg and Tosk).
- The change of *l̥* into *ð* (some Gheg and Lab varieties) and into *ɣ* (some Arbresh dialects).
- Some eastern northern Gheg dialects lack the phoneme *n'* (orthographic *nj*; they have *ni* or *n* instead). In many dialects (both Gheg and Tosk) as well as in SA *n'* changes to *j* in some positions, defined in either phonetic or morphological terms.
- In most Gheg dialects and in a few Tosk dialects there is a special phoneme *ŋ*, which originated from the cluster *ng*.
- The Tosk rhotacism (the change *-n- > -r-*: SA, Tosk *bëra* – Gheg *bāna* ‘I made (aor.)’). Rhotacism applied to both inherited Albanian words and Latin borrowings (SA, Tosk. *rërë* ‘sand’ < Lat. *arēna*). By contrast, only two early Slavic borrowings in Tosk demonstrate the traces of the rhotacism. These facts make it possible to date this change (and the very moment of Albanian dialectal split) to the last centuries of the first millennium AD. The rhotacism is not a positionally conditioned restriction anymore in SA (nor in the Tosk dialects), which has many words with an intervocalic *-n-*. Along with later borrowings, there are inherited words with *-n-*, as well as Latin borrowings among these words; in these words *-n-* originated from various consonant clusters (see below) and had the form of a geminated or long *n*: at the moment of the beginning of rhotacism.

- There is an interdialectal correspondence between *l* in Cham and some Arbresh and Arvanitika and other Tosk diaspora dialects and *j* in most of the Tosk and all of the Gheg dialects as well as in SA that emerged as a result of the development of some clusters with \**l*: \**ly* > *llj* (Cham *milë* ‘thousand’: SA *mijë*); \**li* > *llj* (in the plural forms: Cham *popul* ‘people pl.’ – SA *popuj*). This correspondence makes it possible to reconstruct the sound *λ* for the PA stage (see Schumacher & Matzinger 2013: 245). It seems, however, that this sound was in a complementary distribution with other laterals (*l* and *l̥*) and can be treated as an allophone rather than an independent phoneme (possibly it continued to function as a cluster *ly* at the end of the PA period).

Thus, the following subsystem of sonorants can be reconstructed for the PA stage:

m	n	n:	nʰ
		l̥	l (λ)
	r	r:	
	j		

### Sonorants: the origin

**TABLE 12.5 PIE SONORANTS**

PIE phonemes	PA and SA reflexes	examples
1. * <i>l</i>	a) PA * <i>l</i> , SA <i>l</i> everywhere except intervocalic position b) PA * <i>l̥</i> , SA <i>l̥</i> /V_V (except the position before the * <i>-i</i> , see above, 2.1.)	SA <i>lagje</i> ‘quarter’ < PIE * <i>log<sup>h</sup>-</i> SA <i>hell</i> ‘spear, spit’ < PIE * <i>skōl-</i>
2. * <i>r</i>	a) PA, SA <i>r</i> /V_V b) PA, SA <i>r</i> :/#_	a) SA ( <i>i</i> ) <i>mirë</i> ‘good’ < PIE * <i>miHr-</i> b) SA <i>rrjedh</i> ‘to flow’ < PIE * <i>h<sub>3</sub>reǵ-</i>
3. * <i>m</i>	PA, SA <i>m</i>	SA <i>motër</i> ‘sister’ < PIE * <i>meh<sub>2</sub>-tr-</i>
4. * <i>n</i>	PA, SA <i>n</i>	SA <i>natë</i> ‘night’ < PIE * <i>nok<sup>w</sup>t-</i>

#### Notes

The SA (or PA) *j* arose from the various sources (diphthongization of *e*, development of /*λ*/, the anti-hiatus sound and so on). On the development of the PIE \**y* in Albanian see below (the section on *obstruents*).

1. Lat. *l* has reflexes with a similar distribution in Albanian.
2. The distribution of the two reflexes of PIE \**r* is unclear. The picture is further complicated by the presence of numerous doublets with *r/rr* alternation among the Latin borrowings in the MA (also in the OA texts). PIE \**VrnV* > SA *r*:; PIE \**#sr*, \**wr* > SA *r*:- (see examples in Schumacher & Matzinger 2013: 248). \**-ri* > *j* (in the plural forms: *bij* ‘sons’ < \**bir-i*); whether an intermediate stage was *λ*/ remains unclear.

Lat. *r* > PA, SA *r/r*: both in anlaut and in intervocalic position (with a tendency to follow the pattern of PIE reflexes).

3. On the reflexes of PIE \**m* in some nominal endings see the section “Morphology”. Lat. *m* > Alb. *m*.
4. Lat. *n* > SA *n*; in Tosk, this sound develops into *-r-* in the intervocalic position (see above on rhotacism). Lat. *gn* > SA *nj*.

PIE intervocalic \**-sn-*, \**-Cn-*, \**-nC-* yielded PA \**-nn-*, which further develops into *-n-* in both Tosk and Gheg (for examples, see Schumacher & Matzinger 2013: 249–250).

*Further development of PA sonorants*

On rhotacism and the development of the sound  $\lambda$  see 2.1.

PA \* $nj$ , \* $ni$  > SA  $nj, j$ : SA *brinj* ‘horns pl.’ < PA *brin-i*; SA *běj*, Gheg. *bānj* < PPA *banyō* (on the change  $nj > j$  see 2.1). PA \*- $rn-$  > SA *-r-*.

On the development of  $n$  in closed auslaut syllables see 1.3.

*Syllabic sonorants*

PIE syllabic \* $m̥$  and \* $n̥$  > PA *a*, SA *a*: SA (*i*) *gjatë* ‘long’ < PIE \* $d̥l̥ŋg^h-t-$ ; SA *shtatë* ‘seven’ < PIE \* $septm̥-$ .

The main reflexes of PIE \* $l̥$  and \* $r̥$  are PA, SA *li* and *ri* respectively (a detailed analysis of the distribution of PIE syllabic sonorants in Albanian can be found in Schumacher & Matzinger 2013: 229–232).

**Obstruents**

SA obstruents

p b	t d	c ɟ	k g
f v	θ ð		
	s z		h
	ʃ ʒ		
	ts dz		
	tʃ dʒ		

*Obstruents: main dialectal differences. Reconstruction of the PA system*

- In some dialects (both Tosk and Gheg) the phoneme *h* is absent; in other dialects (mainly Tosk) it is absent in certain phonetic positions. In many Gheg dialects *h* alternates in some positions with *f*. It should be noted that SA *h* often has a non-etymological character. On SA *h* < PIE \* $sk$ , \* $s_k$ , see below.
- The palatal phonemes *c, ɟ* (orthographic *q, g*) manifest in most of the Gheg dialects as affricates (and in some places as fricatives); in some eastern northern Gheg dialects (in Kosovo) they merged with the affricates *tʃ, dʒ* (orthographic *ç, xh*). In the majority of Albanian dialects the palatals *c, ɟ* are the result of the merger of PA \* $c$ , \* $ɟ$  and the PA clusters *kl, gl* (they were registered in this form in some Old Albanian written texts). In some dialects, however, the latter clusters either remain intact (Cham *klumësht* ‘milk’: SA *qumeshtë*; Cham *gluhë* ‘language’: SA *gjuhë*) or yielded other reflexes (western-northern Gheg *ky, gy*; eastern-northern Gheg *k, g*).
- In some southern Gheg dialects the phonemes *θ* and *ð* are absent (they changed into *x* and *ɣ* respectively).
- In the majority of Gheg dialects and some Tosk dialects the consonant clusters *mb, nd* were simplified, respectively, into *m, n* (sometimes into *b, d*).
- Devoicing of voiced obstruents in the word-final position is characteristic of an absolute majority of Gheg dialects and many Tosk dialects (on the potentially common-Albanian character of this phenomenon see Schumacher & Matzinger 2013: 267–270).

Some Albanian obstruents are the relatively late products of inner Albanian developments (various kinds of sandhi, voicing in the position near sonorants and voiced obstruents,

spontaneous consonant strengthening) supported by borrowings. Such are the phonemes *ts*, *dz*, *tʃ*, *dʒ*, *ʒ*. It seems that these phonemes were formed in the PA period (except for the probably more recent *ʒ* and *dʒ*).

Thus, despite interdialectal differences, we should reconstruct the PA obstruent system as nearly identical with the SA system. The only difference is the possible absence of the phonemes *ʒ* and *dʒ* in the PA system. Besides, it can be supposed that the voiced labial spirant (> SA *v*) was bilabial (*β*) in the PA period.

### Obstruents: the origin

#### Occlusives

PIE Media Aspirata merged with Media before the beginning of the PPA period.

TABLE 12.6 PIE OBSTRUENTS

PIE phonemes	PA and SA reflexes	examples
1. *p	PA, SA <i>p</i>	SA ( <i>i</i> ) <i>plotë</i> ‘full’ < PIE *pleh <sub>1</sub> -t-
2. *b <sup>h</sup> , *b	PA, SA <i>b</i>	SA ( <i>i</i> ) <i>bardhë</i> ‘white’ < PIE *b <sup>h</sup> orǵ-
3. *t	PA, SA <i>t</i>	SA <i>tre</i> ‘three’ < PIE *treyes
4. *d, *d <sup>h</sup>	a) PA, SA <i>d</i> in the initial position and in the inlaut after <i>n</i> and before <i>r</i> b) PA, SA - <i>ð</i> - in the inlaut elsewhere	a) SA <i>djeg</i> ‘to burn’ < PIE *d <sup>h</sup> eg <sup>wh</sup> - b) SA <i>hedh</i> ‘to throw’ < PIE *skewd-
5. *k	EPA *ts > Late PA *θ > SA <i>θ</i>	SA ( <i>i</i> ) <i>thellë</i> ‘deep’ < PIE *kel-
6. *ǵ, *ǵ <sup>h</sup>	EPA *dz > Late PA *ð > SA <i>ð</i>	SA <i>dhëmb</i> , Gheg <i>dhâmb</i> ‘tooth’ < PIE *ǵomb <sup>h</sup> -
7. *k	PA, SA <i>k</i>	SA <i>kam</i> ‘to have’ < PIE *keh <sub>2</sub> p-
8. *g, *g <sup>h</sup>	PA, SA <i>g</i>	SA <i>gardh</i> ‘fence’ < PIE *g <sup>h</sup> ord <sup>h</sup> -
9. *k <sup>w</sup>	a) EPA *tʃ > Late PA *s > SA <i>s</i> before front vowels b) PA, SA <i>k</i> elsewhere	a) SA <i>sjell</i> ‘to bring’ < PIE *k <sup>w</sup> elh <sub>1</sub> - b) SA <i>pjek</i> ‘to bake’ < PIE *pek <sup>w</sup> -
10. *g <sup>w</sup> , *g <sup>wh</sup>	a) EPA *dʒw > Late PA *z > SA <i>z</i> before front vowels b) PA, SA <i>g</i> elsewhere	a) SA <i>zorrë</i> ‘gut’ < PIE *g <sup>wh</sup> ern- b) SA <i>djeg</i> ‘to burn’ < PIE *d <sup>h</sup> eg <sup>wh</sup> -

#### Notes

2. A special problem is the reflexes of PPA \*b (as well as other voiced occlusives, see below) in the intervocalic position. On the one hand, there is a change from PPA \*-b- to PA, SA -b- (SA *grabë* ‘erosion, hollow’ < PIE \*g<sup>h</sup>rob<sup>h</sup>-ā); on the other hand, there is also a change from PPA \*-b- to PA, SA #- (SA *det* ‘sea’ < PIE \*d<sup>h</sup>ewb-et-; see another opinion in Schumacher & Matzinger 2013: 233). The conditions of this distribution are not clear (see Çabej 1976: 241–242, Orel 2000: 78–80). The loss of the intervocalic -b- is also attested in Latin borrowings (the reflex of the Lat. *b* in other positions is SA *b*), at least in those borrowings that belong to the older layer: SA *kalë*, Gheg *kâl* ‘horse’ < Lat. *caballum*.
- 4b) In some of the PIE lexemes \*-ð- has fallen out: SA *be* ‘oath’ < \*b<sup>h</sup>oyd<sup>h</sup>-ā. Lat. *d* yielded PA, SA -*ð*- in intervocalic position and after *r*; sometimes \*-ð- (< Lat. -*d*-) disappears. The weakening of \*d (> ð) in the sequence \*-rd continued during the PA period and even later.
- 5–10) Albanian belongs to the few IE languages that retain different reflexes for the three series of PIE gutturals.

- 5–6) PIE \*k > PA, SA *k*, PIE \*ǵ, \*ǵ<sup>h</sup> > PA, SA *g* before resonants: SA *quaj*, Buzuk *kluoj* ‘to call, to give a name’ < PIE \*kluw-.
- 6) Sometimes in the initial position both PIE \*ǵ and \*ǵ<sup>h</sup> yield SA *d*: SA *dorë*; PIE \*ǵ<sup>h</sup>esr-. The conditions of the distribution of the reflexes *d/ð* are unknown. The idea that in anlaut the reflexes of PIE \*ǵ- (> SA *ð*-) and PIE \*ǵ<sup>h</sup>- (> SA *d*-) differ (see Orel 1998: XX; B. Demiraj 1997: 63) does not have convincing etymological evidence.
- PIE \*ǵ<sup>h</sup>w > EPA \*d<sub>3</sub>w > Late PA *z* > SA *z*: SA *zë*, Gheg. *zã* (< *zãn*) < PIE \*ǵ<sup>h</sup>weno- (there are no reliable examples for a parallel development of PIE \*kw and \*ǵw).
- 7–8) With further palatalization into *c, j* before front vowels; see below.
- 9–10) With further palatalization into *c, j* before front vowels; see below.

TABLE 12.7 PIE \*S, \*Y, \*W

PIE phonemes	PA and SA reflexes	examples
1 *s	a) EPA *ž > PA, SA <i>j/##_V</i> b) EPA *š > PA, SA <i>ʃ/##_V</i> c) EPA *ś > PA, SA <i>ʃ/_V(+front)</i> d) EPA *h > PA, SA Ø/V_V(+back) e) EPA *h > PA, SA Ø/_#	a) SA <i>gjak</i> ‘blood’ < PIE *sok <sup>w</sup> - b) SA <i>shi</i> ‘rain’ < PIE *suH- c) SA <i>gǵysh</i> ‘grandfather’ < PIE *suh-s-io- d) SA <i>neve</i> ‘we (gen.)’ < PIE *nōsōm e) SA -# ‘nom. m. sg.’ < PIE *-os
2. *y	a) EPA *ž (?) > PA, SA <i>j/##_V</i> b) PA, SA Ø/V_V	a) SA <i>gjem</i> ‘bridle’ < PIE *yom- b) SA <i>tre</i> ‘three’ < PIE *treys
3. *w	a) PA *w > SA <i>v/##_V</i> b) PA, SA Ø/V_V	a) SA <i>vesh</i> ‘to put on (clothes)’ < PIE *wes- b) SA <i>ve</i> ‘widow’ < PIE *wid <sup>h</sup> ew-

## Notes

1. There is no consensus with respect to the distribution and even the reflexes of PIE \*s in Albanian. It is not known exactly what determines the distribution of *j* and *ʃ* in the initial position. Two other reflexes of PIE \*s (SA *h* for anlaut and inlaut and SA *θ* for anlaut) are now rejected by some researchers (see e.g. Schumacher & Matzinger 2013: 258–265). On some clusters with \*s see below. Lat. \*s > PA, SA *ʃ*, the same outcome is also found in the older layer of Slavic borrowings (in most Slavic borrowings *s* reflects as SA *s*). For some chronological considerations related to the development of PIE \*s, see below.
2. See Schumacher & Matzinger 2013: 252–254; for other opinions see Čabej 1976; de Vaan *in print*. The distribution of the reflexes of Lat. /y/ is, generally speaking, similar to that of PIE \*y (sometimes Lat. intervocalic /y/ > SA *j*).
3. The distribution of the reflexes of Lat. /w/ is, generally speaking, similar to that of PIE \*w.

## Laryngeals

On the reflexes of the PIE laryngeals that were adjacent to vowels, see 1.2. When adjacent to consonants, laryngeals disappeared. For the possible influence of the laryngeals on Albanian consonants see B. Demiraj 1997: 60–61. For the influence of the laryngeals on the neighboring syllabic sonorants see Schumacher & Matzinger 2013: 230–231. For the development PIE \*Hr- > PA, SA *r*- see B. Demiraj 1997: 70–71.

## Reflexes of some PIE consonant clusters in Albanian

- PIE \*sk/sk/sk<sup>w</sup> > PA, SA *h*: SA *hedh* ‘to throw’ < PIE \*skewd-.
- PIE \*sp > PA, SA *ʃ/##\_*: SA *farë* < PIE \*spor-. Another source of PA, SA *ʃ* is borrowings.

PIE \*st > PA, SA *ft*: SA *shteg* ‘path, road’ < PIE \*stoyg<sup>h</sup>-.

PIE \*zd > PA, SA *ḍ*: SA *pidh* ‘female pudenda’ < PIE \*peysd<sup>(h)</sup>-. Lat. *sc*, *sp*, *st* give, respectively, SA *sk*, *sp*, *st*.

Clusters of the type “dental + y,” “guttural (< PIE velars and labiovelars) + y” developed into EPA \*tʃ, \*dʒw, PA, SA *s*, *z*, depending on whether the occlusive was voiced or voiceless.

For the development of some other consonant clusters see B. Demiraj 1997: 54–68, Matzinger 2006: 68–81, Schumacher & Matzinger 2013: 243–244.

### *Chronology of Albanian palatalizations*

As mentioned above, there were in Albanian different processes of palatalization of PIE velars (sometimes developed further into assibilation): (a) “satem”-assibilation of PIE palatals gave rise to interdental spirants *θ* and *ḍ*; (b) assibilation of PIE labiovelars (and of the groups “palatal + w” but not of the “pure” velars) before the front vowels resulted in *s*, *z*; and (c) palatalization of \*k, \*g (from both PIE velars and labiovelars), also before front vowels, resulted in *c* and *j*.

- a) As for the satem-palatalization, a common opinion is that there was an intermediate stage of the development, namely the affricates of the type *ts*, *dz*. It is not known exactly when the process of assibilation took place, but it is possible to assume that it appeared in the first half of the PA period (the reason for this hypothesis is that various Latin palatalized consonantal sequences never reflect as SA *θ* or *ḍ*). It is not completely clear when the result of the development of the PIE voiced palatal merged with \*d < PIE \*d<sup>(h)</sup> (and whether this process could have influenced the distribution of the reflexes *d/ḍ* that go back to PIE \*ǵ<sup>(h)</sup>).
- b) The assibilation of PIE labiovelars before the front vowels had to take place also relatively early (before the change PIE \*ē > PPA *ā*; see SA *zorrë* ‘gut’ < PIE \*ǵ<sup>wh</sup>ēr̥n-). The reflexes of PIE labiovelars merged in SA with the results of the development of the PIE groups \*ty, \*dy, \*k<sup>(w)</sup>y, \*g<sup>(w)</sup>y as well as Lat. *tj*, *dj*. This fact as well as the reflection of the *s* of Latin and early Slavic borrowings as SA *sh* indicates convincingly that in the EPA period the reflexes of labiovelars must have been represented a kind of affricates, and the change into the *s* had to take place at the very end of the PA period.
- c) The process of palatalization that resulted in *c* and *j* sounds belongs, on the contrary, entirely to the PA period. This change took place in the positions before *i*, *y* and *e* (including the *je* and *ja* reflexes of PIE \*e). This wave of palatalization should have begun *before* the change in the unstressed vowels (see SA *lplec* ‘old men (pl.)’ < \*plaki; SA *virgjër* ‘virgin’ < Lat. *virginem*) and continued to be at work after the umlaut processes (that is, *e* < \*a did trigger the palatalization, which explains e.g. this pattern of palatalization of PIE labiovelars: OA *qell* ‘to bring’ < EPA \*kal-it < PIE \*k<sup>w</sup>olh<sub>1</sub>-). Lat. *c*, *g* before front vowels as well as Lat. clusters *cy*, *gy* also give SA *c*, *j*. In the later borrowings, *k*, *g* are reflected as *c*, *j* rather seldom and can be explained as a result of analogical processes. In the PA period the palatalized reflex of \*g merged with reflexes of the PIE \*s- and \*y- (which may have led to the phonologization of *j*). During the OA period *c*, *j* merged in the majority of Albanian dialects with the reflexes of the PA clusters \*kl, \*gl (see 3.1 above).



## Stress

The stress in modern Albanian is dynamic. It shows a clear tendency to be columnal and to fall in the vast majority of cases on the stem. In effect, the morphological forms of words typically are not differentiated by the position of the stressed syllable. The only exception is a small group of nouns and adjectives that have different positions of the stress in the singular and plural forms. These include some nouns with the plural suffixes *-inj*, *-ij*, *-enj* (SA sg. *dr'apër* 'banner' – pl. *drapër'inj*, SA sg. *l'umë* 'river' – pl. *lum'enj*), SA sg. *dh'ëndër* 'groom, son in law' – pl. *dhënd'urë*. Two adjectives (SA sg. m. (*i*) *madh* 'big' – pl. m. (*të*) *mëdh'enj*, SA sg. m. (*i*) *keq* 'bad' – pl. m. (*të*) *këq'ij*) have an additional vocalic alternation. The same rightward shift of the stress in plural forms is also observed in Turkish borrowings with the plural suffix *-llarë/-lerë* (SA sg. *ag'a* 'aga' – pl. *agall'arë*); however, in these words the pattern is explained by the parallel borrowing of both singular and plural forms (the very suffix *-llarë/-lerë* is used only with Turkish loanwords). Two nouns have a leftward stress shift in the plural: SA sg. *njer'i* 'man' – pl. *nj'erëz*, SA sg. *kall'i* 'ear' – pl. *k'allëza*. It can be noticed that in all these cases there is a non-productive pattern with clearly different singular and plural stems. The majority of derivational suffixes in Albanian are stressed.

As for the position of the stress within the word, it is relatively free. There is a tendency to place the stress on the penultimate syllable; however, there are multiple exceptions to this pattern. For example, there are many uninflected words with the last syllable stressed (*edh'e* 'still' and so on). As for inflected words, the majority have forms with the penultimate syllable stressed (on the rules of the SA accent see Buchholz & Fiedler 1987: 53).

Albanian dialects have relatively few differences with respect to their stress patterns. The tendency to stress the penultimate syllable is strongest in Gheg dialects: there are words (nouns and pronouns) with stress on the penultimate syllable in Gheg, whereas in Tosk counterparts the stress is on the last syllable (SA, Tosk *njer'i* 'man' – Gheg *nj'eri*). There is also a tendency in Gheg to shift the stress towards the penultimate position in Turkish loans (SA, Tosk *par'a* – Gheg *p'are* 'money' < Turk. *par'a*). Dialectal variation is also observed in the realization of stress in diphthongs *ue*, *ye*, *ie* in various phonetic contexts (see "Vowels").

Generally, the stress pattern in old written texts is close to the modern system. The main difference is a weaker degree of reduction in unaccented vowels, in particular, the vowels in word-final syllables. As a result there are a number of words in these texts with their prepenultimate syllable stressed that correspond to SA words with the penultimate syllable stressed (e.g. old Gheg *gj'arpënë* 'snake' – SA *gj'arpër*). For possible differences in the stress position in verbal non-active forms see "Morphology".

Thus, at the end of the PA period the stress patterns must have been very near to the modern ones, the difference mainly amounting to a greater variability in the position of stress in the word, which is accounted for by a lesser degree of vowel reduction in PA.

It is well known that Latin borrowings (as well as later loans; see, however, above on Turkish borrowings) maintained their accent patterns in Albanian. It can be hypothesized, then, that immediately before the beginning of the Latin-Albanian contacts (that is, at the end of the PPA period) Albanian stress patterns didn't differ essentially from the late PA ones. For vowel reduction, which was accounted for by the dynamic character of Albanian stress and changed the shape of Albanian words, see above ("Vowels").

It is not possible to define the time limits and the exact rules of the transition from the late IE stress system to the Albanian one. The only thing that is certain is that Albanian somehow abandoned before the beginning of the PA period both the IE final accent and accentual mobility within nominal and verbal paradigms (Schumacher & Matzinger 2013: 208–209).

## Morphonological alternations

Albanian is a language in which morphonological alternations play a relatively important role in the expression of morphological oppositions. Most of these alternations reflect the phonetic changes in the PA period (on umlaut, diphthongization, change  $e > i$ , see “Vowels”; on palatalization, see “Obstruents”; on loss of the final  $-n$ , development of  $*-rn-$ , see “Resonants”). The old PIE alternations, such as first of all ablaut, left but some minor traces in the Albanian verb system. For more information on the realization of the main morphonological alternations, see “Morphology”.

## MORPHOLOGY

0. Albanian is a fusional language that combines synthetic and analytic features. Synthetic features are only partly inherited from the PIE state, while a considerable part of the flexion, especially in the nominal system, is the result of the processes of the new synthesis that took place in the PA and even in the OA periods. The level of analyticity is, however, rather high. It is fairly typical of the numerous analytic constructions to make broad use of clitics, in both the nominal and verbal domains. Albanian has some grammatical traits that are rather unusual from an areal-typological perspective: a morphological optative and the analytic double case agreement; see below. At the same time, many grammatical features, both morphological and syntactical, identify Albanian as a language that belongs both to the Balkan Sprachbund and to the languages of the Standard Average European type. The morphological part of this chapter is based, first of all, on Buchholz and Fiedler 1987 (for SA), Gjinari et al. 2007 (for dialectological material), and Sh. Demiraj 1986, Matzinger 2006, Schumacher and Matzinger 2013, Fiedler 2004, Klingenschmitt 1981, 1994, and Orel 2000 (for historical and comparative interpretation).

## Noun

### *The SA noun system*

*Grammatical categories of Albanian nouns.* The noun in SA is characterized by the categories of gender, number, case and definiteness. The *gender* of the noun is defined by the agreement features of the dependent elements (determiners, agreeing adjectives and so on) and has three main values: masculine, feminine and neuter. For animate nouns the gender (masculine or feminine) is mostly semantically determined. The neuter encompasses derived deverbal and deadjectival nouns with abstract meaning as well as a handful of old concrete nouns (all of these concrete nouns also have masculine or feminine variants that are more common in SA). Feminine and neuter nouns don't differ in the plural by their agreement properties. Besides, there is a fourth agreement class that includes inanimate nouns that agree according to the masculine model in the singular and according to the feminine model in the plural (this is called “ambigender” in the Albanian grammatical tradition). On the inflectional morphology of Albanian gender see below.

The category of *number* has two values: singular and plural. The rules of the derivation of plural forms are extremely complex and opaque (see below).

Albanian has the following *cases*: nominative, accusative, genitive-dative and ablative; the ablative is distinguished from the genitive-dative only by one optional form in the noun paradigm as well as in certain pronominal paradigms. Genitives and datives, which are traditionally differentiated in Albanian grammars, actually differ syntactically. The form that is used in the possessive constructions (= genitive) is always preceded

by a special prepositive clitical formant (“article” in the Albanian grammatical tradition), which agrees with the nominal head of the construction: *djal-i i fshatar-it* boy-NOM.M.SG.DEF. ART.NOM.M.SG. peasant-GEN./DAT.M.SG.DEF. (on the forms of this “article” see “Adjective” below). The form that is used in the syntactic position of the indirect object and sometimes adverbially (= dative) requires the use of the pronominal clitic before the verb form (see “Syntax” below).

*Definiteness* is expressed in Albanian by the means of the postpositive article, which is agglutinated to the singular and plural noun stems. Thus, there are two sets of the nominal paradigms in Albanian: the definite and the indefinite ones. Besides, there is a prepositive indefinite article (homonymous to the numeral ‘one’). Thus, the category of definiteness/indefiniteness is best analyzed as having three values in Albanian.

*Declension types of Albanian nouns.* The Albanian noun has four declension classes in the singular; see Table 12.8.

**TABLE 12.8 NOUN DECLENSION IN SINGULAR**

	I (masculine)		II (masculine) <sup>2</sup>		III (feminine)		IV (neuter)	
	indef.	def.	indef.	def.	indef.	def.	indef.	def.
nom.	<i>mur</i> <sup>1</sup>	<i>muri</i>	<i>mik</i>	<i>miku</i>	<i>vajzë</i>	<i>vajza</i>	<i>të<sup>s</sup> folur</i>	<i>të folurit</i> <sup>6</sup>
acc.	<i>mur</i>	<i>murin</i>	<i>mik</i>	<i>mikun</i>	<i>vajzë</i>	<i>vajzën</i> <sup>4</sup>	<i>të folur</i>	<i>të folurit</i> <sup>6</sup>
gen.-dat./abl.	<i>muri</i>	<i>murit</i>	<i>miku</i>	<i>mikut</i>	<i>vajzë</i> <sup>3</sup>	<i>vajzës</i> <sup>4</sup>	<i>të foluri</i>	<i>të folurit</i>

Notes: *mur* ‘wall’, *mik* ‘friend’, *vajzë* ‘girl’, *të folur* ‘speech’.

All words and word forms in the sections on SA morphology are from Standard Albanian, unless indicated otherwise.

Each class can be further divided into minor subclasses, which would differ by the alternations of the stem in various case forms and, in the cases of classes III and IV, by the phonemic and/or orthographic shapes of the endings.

The correlation between the declension class and gender isn’t straightforward (except for the IV class, which is exclusively neuter). The exceptions, however, aren’t numerous (these are mostly proper nouns). Besides, the numerous ambigender names belong to the I and II subclasses.

- 1 Although there is a homonymous nominative and accusative form in the indefinite declension, these paradigmatic cells are differentiated by the form of agreeing elements.
- 2 The Class II includes all the nouns whose stem ends in *-k*, *-g*, *-h*, a portion of the masculine nouns that end with a stressed vowel, as well as the word *krye* ‘head’.
- 3 In some subclasses the ending is *-je*.
- 4 In some subclasses the endings are *-në*, *-së*.
- 5 Deverbal and deadjectival nouns retain the prepositive “article” of the adjectives (see “Adjective” below).
- 6 In some subclasses the ending is *-(ë)t*, *-të*.

The plural forms a single declension class that is divided into subclasses that differ according to the phonetic/orthographic shapes of endings depending on the plural stem auslaut. It is noteworthy that this auslaut typically is not determined by the gender of the noun. The general schema of the plural declension is as follows:

TABLE 12.9 NOUN DECLENSION IN PLURAL

	indefinite	definite
nom./acc. <sup>1</sup>	<i>mure</i>	<i>muret</i>
gen.-dat.	<i>mureve</i> <sup>2</sup>	<i>mureve</i>
abl.	<i>mureve/muresh</i> <sup>3</sup>	<i>mureve</i>

## Notes

- 1 The nominative and accusative forms differ neither morphologically nor by the shape of the agreeing elements.
- 2 In the plural, there is no opposition between indefinite and definite forms for the genitive-dative case.
- 3 The indefinite ablative ending has two allomorphs. Their sphere of use is determined (with many exceptions) by the syntactic position of the ablative word forms.

*Derivation of nominal plural forms.* Only for some nouns can the plural form be predicted based on the singular one (for other nouns the plural is idiosyncratic).

For masculine nouns the plural form is derived by means of several consonantal (*k/q, g/gj, ll/j, n/nj, r/j*) and vocalic (*a/e*) alternations, as well as by means of more or less productive suffixes (*-e, -ë, -a, -ra, -inj, -llarë/lerë*). Some masculine nouns have identical stems in the singular and in the plural. Plural suffixes for feminine nouns are *-a* and *-ra*, but a considerable portion of feminine nouns do not differentiate singular and plural stems. There are a few nouns, both masculine and feminine, that have irregular plural forms. Most of the old neuter nouns have the plural suffix *-ra*. Some SA nouns have doublet plural forms.

### Basic dialectal differences in the nominal system

- 1) An additional case – the locative – is preserved in both old texts and in the dialects (northern Tosk and southern and central Gheg). This case is used after some prepositions with the locative meaning (*në mal-t* ‘on the mountain’, *në male-t* ‘on the mountains’) and doesn’t differ formally from the accusative in the indefinite form.
- 2) In SA, the ablative is the case that tends to merge with the genitive-dative. A further step of this development is observed in some diaspora dialects (Arvanitica, Ukraine, Bulgaria), where the ending *-sh* disappeared completely from the case system. By contrast, a consistent formal opposition of the ablative and genitive-dative is observed in the old written texts. Some remnants of this situation are preserved in the western northern Gheg dialects, in which there is an opposition of the ablative and genitive-dative forms for feminine definite nouns (abl f. def. *delet* ‘sheep’ – gen.-dat f. def. *delesë*). A dedicated ending *-shit* for the ablative plural definite form existed in the old written texts.
- 3) There are fluctuations in gender assignment of individual words. Thus, the number of underived neuter nouns in the earliest written texts as well as in some dialects is bigger than it is in SA. There are also gender fluctuations between masculine and feminine genders. In the first written texts and in some Gheg dialects there is a special subclass of feminine nouns of Latin origin that correspond to SA masculine nouns: SA (and majority of dialects) *shëndet* m. ‘health’ – OA *shëndet* f.

- 4) The fluctuations in the sphere of definiteness are manifested in the varying degrees that the opposition “definite : indefinite” is reflected in the nominal paradigm. Thus, in the old texts and some dialects there was a special definite ending (*-vet*) for the genitive-dative case in the plural, which is absent in SA. By contrast, in some dialects the weakening of the “definite : indefinite” opposition has gone further than it did in SA.
- 5) There are many dialectal variations in the means of plural formation. For example, the plural suffix *-a* is systematically used in Gheg dialects, whereas the suffix *-ra* is frequently found in northern Tosk dialects.
- 6) There are differences in the form of individual endings. The following cases are especially relevant from the historical point of view:
  - a) Many dialects as well as the old texts show the masculine definite accusative ending *-në* without the “connecting” vowel *i/u*: SA *mikun* ‘friend (dat.)’ – dial. *miknë*.
  - b) In Buzuk’s book, the masculine definite genitive-dative form is formed not only with the endings *-it/-ut* but also with the ending *-(ë)t*, which is attached to those nouns whose stem ends in *-ë*: Buzuk (*i*) *djalët* ‘boy (gen.)’. In some dialects and in the old texts, the masculine definite genitive-dative forms of some kinship terms are formed with the help of endings without “connecting” vowels: Buzuku *d birt* ‘son (gen.)’.
  - c) In the old texts as well as in some dialects (both Tosk and Gheg) the genitive-dative plural ending has the form *-e* after the stems that end in consonants (*pleq-e* ‘old man (gen.-dat. pl.)’) and *-ve* after the stems that end in vowels (*plaka-ve* ‘old woman (gen.-dat. pl.)’).
- 7) There are, finally, morphological distinctions that are due to phonetic dialectal differences: SA and Tosk plural endings *-ër, -ra* – Gheg. *-(ë)n, -na*. In Gheg dialects there is a compensatory apocope as a consequence of the loss of the final *-ë*: Gheg *pla:k* ‘old woman’ – SA *plakë*.

### Nominal system: the origin

Albanian has radically rebuilt the PIE noun system: its declension system has been regularized, and the “old” PIE endings have been for the most part lost. The details of this process cannot be reconstructed, but it can be supposed that it began in the PPA period (the time of an essential rebuilding of the PIE stress system in Albanian) and continued at least till the end of the PA period (the time when the process of unstressed vowel reduction was still under way).

In many cases, this radical rebuilding makes it impossible to establish a reliable historical source of Albanian flexions. Many nominal flexions are reconstructed on the basis of our general ideas about the development of the PIE grammatical system (this is also true for the history of the verbal system). As a consequence, in many cases there are competing etymological hypotheses; below I will provide only the most reliable of these hypotheses. In those instances when there are competing proposals, the relevant bibliographical references are given, although in most cases I provide references only for modern overviews rather than for those original studies where these respective solutions were proposed.

#### *The indefinite declension*

In this section and below, when discussing various historical explanations I will employ PIE notations in the form that was used by the respective authors.

## Masculine

The “masculine” declension type (I and II declension types in Table 12.8) continues the PIE *o*-declension (with the *-yo-* subtype). In Albanian, these types encompass masculine nouns that used to belong to other declensions in PIE (that is, the consonant stems, *u*-stems, *i*-stems and so on). Some nouns that belonged to the PIE *i*-declension type manifest traces of the PA umlaut processes, which indicates that the final stage of the rebuilding of the declension system happened in the PA period (SA *elb* ‘barley’ < PIE \**h*<sub>2</sub>*olb*<sup>h</sup>*i-*).

The flexion of the indefinite declension is deeply decayed due to reduction that continued to happen in the PA period (these processes affected Latin borrowings; see “Phonology”). The following endings for the *o*-declension can be reconstructed:

*nom. sg.* SA *-ë*, *-Ø* < PIE \**-os*. The double reflex *ë/Ø* is conditioned by the phonetic shape of Albanian words, although the distribution isn’t completely clear. Orel (2000: 233) explains the *-ë* ending as a reflex of PIE oxytonic stress, but this explanation isn’t supported by the majority of specialists.

*acc. sg.* SA *-ë*, *-Ø* < PIE \**-om* (see also below in the section devoted to the definite declension).

*gen.-dat./abl. sg.* SA *-i*, *-u*. (a) < PIE \**-osyo* (the genitive ending of *o*-declension) (Klingenschmitt 1994, Matzinger 2006: 98, Schumacher & Matzinger 2013). The *i/u* alternation is explained as the result of the development of unstressed *o* under different phonetic conditions (through an intermediate stage \**ü*). (b) < PIE \**-ōy* (the dative ending of the *o*-declension) or \**-oy/-ey* (the locative ending of the *o*-declension; see Orel 2000: 234); see the overview in Sh. Demiraj 1986: 274–282; *-u* is a secondary, phonetically conditioned variant.

## Indefinite

The *nom. pl.* endings demonstrate significant variability:

- a) The fact that these forms are characterized by various kinds of consonant alternations (palatalizations) and vocalic alternations (umlaut) (see “Phonology” above) enables us to trace them back to originally pronominal ending: SA *-Ø* < PA \**-i* (which caused palatalization and umlaut) < PIE \**-oy*. The morphological alternations that emerged in the *nom. pl.* form later got extended analogically into other members of the plural paradigm.
- b) SA *-ë* has different etymological explanations: (a) *-ë* < PIE \**-ōs* (Sh. Demiraj 1986: 231 with some reservations; Orel 2000: 235); (b) *-ë* < PIE \**-es* (initially a consonant stem ending; Matzinger 2006: 102); and (c) *-ë* < PIE \**-ā* (the unstressed ending of the PIE collective; Desnickaja 1976: 37–38).
- c) SA *-e*. (a) < PIE \**-ōs*, \**-oy* in stressed position (Orel 2000: 235); (b) < PIE \**-yā* (the ending of the PIE collective; Desnickaja 1976: 37–38); and (c) < PIE \**-ew-es*, which is a development of the PIE *u*-stems (Klingenschmitt 1994: 25, Matzinger 2006: 102). Extension by analogy from the feminine declension is also plausible.
- d) SA *-a*. (a) This ending arose due to the analogical influence of feminine nouns (Orel 2000: 236; Topalli 2011: 242); (b) < PIE \**-ā* (the stressed ending of the PIE collective; Desnickaja 1976: 38–39); (c) contamination of the endings *-ë* and *-e* (Pedersen 1895: 10, Matzinger 2006: 102; cf. similar contractions in pronominal clitics, see below).

The plural suffixes *-ra/-na*, *-ër/-(ë)n*, *-inj* and some others are in all probability the result of the reanalysis of the old singular *-n* stems and their further concatenation with various plural endings.

The *acc. pl.* ending is always identical with the *nom. pl.* ending. Somewhat conventionally, it is possible to reconstruct the PIE \**-ons* > SA *-Ø* (probably, sometimes *-ë*)

development and later analogical extensions of various types of nom. pl. endings into the acc. pl. form (Orel 2000: 236).

The *gen.-dat. pl. -e* < PIE \*-ōm. The *v* component in the ending *-ve* has emerged as a means to avoid hiatus and was extended analogically from those plural stems that terminate in vowels onto stems of other types.

The *abl. pl. -sh* < PIE \*-oy-su. Unstressed *-u* developed in PA in *-i*; see OA definite abl. pl. *-shit* (see below, Schumacher & Matzinger 2013: 263).

### Feminine

Feminine declension types also underwent a very serious uniformation, mostly using the PIE \*-ā (\*-eh<sub>2</sub>) declension type as a model. Other types of PIE feminine declensions were included in the *-a* declension.

*nom. sg. endings* The most widespread feminine stem-final *-ē* is the reflex of PIE \*-eh<sub>2</sub>.

SA feminine nouns with unstressed final *-e* are interpreted as reflecting the PIE *-ih<sub>2</sub>/-yeh<sub>2</sub>*-stems (“*devī*-type”: SA *hije* ‘shadow’ < PIE \*ski-yeh<sub>2</sub>); see Matzinger 2006: 96.

*acc. sg.* SA *-ē, (-e)* < PIE \*-eh<sub>2</sub>m (see also below, the section devoted to the definite declension).

*gen.-dat./abl. sg.* SA *-e* can be traced back to alternative sources: (a) the PIE dative ending \*-eh<sub>2</sub>ey (Orel 2000: 238) or the PIE locative ending \*-eh<sub>2</sub>i (Sh. Demiraj 1986: 285); and (b) the analogical development from the masculine *gen.-dat./abl. sg. ending*, so we may reconstruct something like \*-ayya < \*-asya (Klingenschmitt 1994: 223, Matzinger 2006: 99–100).

Feminine *nom. pl. endings* are characterized by great variability, although to a lesser degree than masculine *nom. pl. endings*:

- a) SA *-ē* < PIE \*-eh<sub>2</sub>es.
- b) SA *-e* < PIE \*-ih<sub>2</sub>es. Desnickaja (1976: 40–41) considers both *sg.* and *pl. feminine -e* forms as the continuation of the PIE collective in \*-yā.
- c) SA *-a.* (a) < PIE stressed \*-ās (\*-eh<sub>2</sub>es) (Orel 2000: 239); and (b) contamination of the endings *-ē* and *-e* (Matzinger 2006: 103).

The *acc. pl. ending* is always identical with the *nom. pl. ending*.

The *gen.-dat. pl. -e* as well as *abl. pl. -sh* are identical to the masculine ending. They result from the development of PIE \*-eh<sub>2</sub>om (or \*-āsōm; see Klingenschmitt 1994: 224) and \*-eh<sub>2</sub>su endings, respectively.

Both in the masculine and feminine gender, the plural endings of oblique cases are agglutinated to the plural stem, which simultaneously functions as the nominative plural form (after some types of stems that end in a consonant, a connecting vowel is used). It should be stressed that the morphological alternations that first developed in the indefinite nominative plural form of masculine nouns later extended into all members of the plural paradigm.

### Neuter

The only ending that is different from the masculine ones in the singular is *nom.-acc.* of the definite conjugation (see below). According to Pedersen (1897: 288), the *nom.-acc.*

sg. indef. SA *-ë/Ø* of “concrete” neuter nouns goes back to PIE *\*-ā* (*\*-eh<sub>2</sub>*) (nom.-acc. pl. of PIE neutra).

On the origin of some irregular plural forms see Sh. Demiraj 1986: 244–248; for the stress shift in some plural types see “Stress” above). See also an interesting example of umlaut in the gen.-dat./abl. sg. of *at* ‘father’: *et(i)* (< *\*ati*).

### *The definite declension*

The definite nominal forms originated from concatenation of indefinite forms with the corresponding case forms of the postpositive definite article, which in its turn evolved from the demonstrative pronoun. The postpositive article is one of the main balkanisms of Albanian that is shared with Balkan Romance and Balkan Slavic. The definite declension must have been formed in the PA period, it is very likely that the time of this process coincided with the period of intensive (proto)Albanian-(proto)Romanian language contacts (in this case Albanian must have been the source language for this shared feature).

There are two main etymological hypotheses about the origin of demonstrative pronouns. According to the first of them, they evolved from an amalgam of three different PIE pronouns, namely, *\*is*, *\*ih<sub>2</sub>* (*\*yā*) for nom. m and f. sg, *\*ki-/\*ky-* for gen./dat./abl. f. sg. and *\*so*, *\*seh<sub>2</sub>*, *\*tod* for the remaining forms (Meyer 1892: 79, Orel 2000: 243). According to the second hypothesis, all Albanian demonstrative pronominal originated from PIE *\*so*, *\*seh<sub>2</sub>*, *\*tod* (Pedersen 1900: 311–315, Matzinger 2006: 109–110).

### Masculine

*nom. sg.* SA *-i*, *-u* < PIE *\*-os* + *\*is* (Meyer 1892: 40, Orel 2000: 246); (b) PIE *\*-os* + *\*so* (Pedersen 1894: 251, Klingenschmitt 1994: 224). Both etymologies assume that the phonetically conditioned distribution of the forms in *-i* and *-u* arose later than these forms as such; in terms of conditions this distribution is the same as the distribution of formally similar endings of the *gen.-dat./abl. sg. indefinite* (see above).

*acc. sg.* *-(i/u)n* < PIE *\*-om* + *\*tom*. The *i/u* evolved under analogical influence from the *gen.-dat. sg.*

*gen.-dat./abl. sg.* SA *-it*, *ut* are the result of the concatenation of the forms of the indefinite declension with the cased forms of the demonstrative pronoun (that is, *\*tosyo*, *\*tōy* or *\*tey* according to alternative etymological hypotheses; see above). The old and dialectal forms in *-ët*, *-t* can be explained by accentological considerations (Matzinger 2006: 98).

*nom. pl.* SA plural stem + *-t(ë)* < PIE *\*toy*.

*acc. pl.* SA plural stem + *-t(ë)* < PIE *\*tons*.

*gen.-dat. pl.* SA *-ve* is identical with the corresponding indefinite ending (see above). The older ending *-(e/ve)-t* could have evolved from PIE *\*tōm*.

In the old SA ending of the abl. pl. *-shit*, *-t* could have evolved from PIE *\*toysu*; on the change unstressed *-u-* > *-i-* see above (“Indefinite Declension”).

### Feminine

*nom. sg.* SA *-a*: (a) < PIE *\*-ā-* + *\*yā* (Orel 2000: 247); (b) < PIE *\*-ā-* + *\*sā* (Klingenschmitt 1994: 223)

*acc. sg.* SA *-(V)n* < PIE *\*(V)m* + *\*tām*

In the *gen.-dat. sg.* SA *-(V)s(ë)-s(ë)* could have evolved from PIE *\*tesyasyo* (according to Matzinger 2006: 99) or *\*kyāy* (according to Orel 2000: 247).



The old and dialectal *abl. sg.* ending *-et* evolved due to analogical influence from the masculine declension type.

From a synchronic point of view singular oblique feminine forms of the definite declension are, unlike masculine forms, merely combinations of the feminine indefinite stem with the case forms of the definite article.

*nom. and acc. pl. SA:* plural stem + *-t* < PIE \* *tās*.

*gen.-dat. pl. SA -ve:* see above (“Masculine”); older *-t* could go back to \**teh<sub>2</sub>om* (or \**tāsōm*; see Matzinger 2006: 100).

*abl. pl. SA -shit:* see above (“Masculine”); *-t* could arise according to the masculine model (Klingenschmitt 1994: 224).

It is possible that feminine plural definite declension in general has been restructured according to the masculine pattern.

### Neuter

*nom.-acc. sg. -(i, ě)t(ĕ)* < PIE \**-ā-* + \**tā* (Pedersen 1897: 288) or PIE \**-od* + \**tod* (see Matzinger 2006: 99).

### Adjective

There are *two main classes of adjectives* in SA, which differ in their morphological features as well as in the number of agreeing categories they possess. The first class (henceforth *articulated adjectives*) comprises those adjectives that are always used with a special prepositive clitic formant (*article* in the Albanian grammatical tradition) that agrees with the head noun in gender, case and number and is formally identical with the “article” of the genitive noun form (see above). No other grammatical formants can be placed between the article and the adjective.

The second class (henceforth *unarticulated adjectives*) includes those adjectives that have no such “article” and agree with the head in gender and number only by means of their endings.

All Albanian adjectives are strictly divided into these two classes.

The first class, articulated adjectives, includes underived adjectives of Indo-European origin as well as Latin and Slavic borrowings, adjectives derived by some suffixes (*-t(ĕ)*, *-mĕ*, *-shĕm*), adjectives with the negative prefix *pa-*, composite lexemes that have an adverb as their first part, and, finally, deverbals (participles used with the prepositive “article”).

The second class, unarticulated adjectives, includes a small group of old words that can be used as either nouns or adjectives (*plak* ‘old man’ and ‘old’, *trim* ‘brave man, warrior’ and ‘brave’, etc.), the majority of derived adjectives (many of them are also homonymous with nouns), the majority of composite adjectives, and later borrowings.

The normal *word order* in Albanian is “noun + adjective”. The reverse order is possible, too, but often has an additional expressive value. The majority of unarticulated adjectives cannot be used in the construction with the inverted order (composite adjectives make an exception). If an adjective (or an adjective-like pronoun or an ordinal numeral) occupies the first position in the noun phrase, it declines as a noun, whereas the noun loses its ability to mark the case in this construction (see the table). This rule according to which it is always the first component of the noun phrase that get case inflection is a characteristic that Albanian shares with Romanian and Bulgaro-Macedonian.

There are no traces of synthetic formants of *degrees of comparison* in Albanian. Both comparative and superlative meanings of adjectives and adverbs are expressed by the

construction *më* (Gheg *ma*) ‘more’ (< PIE \*meh<sub>2</sub>-is) + adjective or adverb. In some constructions nouns can also have degrees of comparison.

*The declination of adjectives*

**TABLE 12.10 ARTICULATED ADJECTIVES (UNMARKED ORDER)**

	singular		plural	
	masculine	feminine	masculine	feminine
nom.	<i>djali i mirë</i>	<i>vajza e mirë</i>	<i>djemtë e* mirë</i>	<i>vajzat e* mira</i>
acc.	<i>djalin e* mirë</i>	<i>vajzën e* mirë</i>	<i>djemtë e* mirë</i>	<i>vajzat e* mira</i>
gen./dat./abl.	<i>djalit të mirë</i>	<i>vajzës së* mirë</i>	<i>djemve të mirë</i>	<i>vajzave të mira</i>

- a) The table reflects the most widespread declination type of articulated adjectives. Besides this type, there are (a) adjectives with suffixes *-shëm*, *-(ë)m*, *-më*, which have a morphological opposition of the forms of the masculine (both sg. and pl.: *i mesëm*, *të mesëm* ‘middle’) and feminine (both sg. and pl.: *e mesme*, *të mesme* ‘middle’) gender; (b) a small number of adjectives that do not change their endings at all, and (c) some irregular adjectives.
- b) The case forms of the “article” in noun phrases with adjectives do not differ from those forms of articles that are used in noun phrases with the genitive (see above).
- c) The forms given in the table are used when the head noun is in the definite form, and the adjective occupies an adjacent position and functions as an attribute (not as a predicate). Whenever any of these conditions is violated, those forms of articles that are marked with the asterisk in the table are replaced by the article form *të* (*vajzën më të mirë* ‘the best girl (acc.)’, and so on).
- d) Neuter nouns in the singular require the article form *e* for the nom./acc. and *të* for oblique cases when the noun phrase has an attributive dependent and the article occupies the position immediately after the definite head noun. Otherwise, the form *të* is used for all cases. In the plural, the adjectives change according to the feminine pattern.

**TABLE 12.11 ARTICULATED ADJECTIVES (INVERTED ORDER)**

	singular		plural	
	masculine	feminine	masculine	feminine
nom.	<i>i miri djalë</i>	<i>e mira vajzë</i>	<i>të mirët djem</i>	<i>të mirat vajza</i>
acc.	<i>të mirin djalë</i>	<i>të mirën vajzë</i>	<i>të mirët djem</i>	<i>të mirat vajza</i>
gen./dat./abl.	<i>të mirit djalë</i>	<i>së mirës vajzë</i>	<i>të mirëve djem</i>	<i>të mirave vajza</i>

Substantivized articulated adjectives decline identically to adjectives in the noun phrases with inverted word order.

*Unarticulated adjectives* agree with the head nouns only in gender and number. They are divided into several inflectional classes. In the singular, the most widespread endings are zero for masculine and *-e* for feminine gender. In the plural, the most typical endings are *-ë* for masculine and *-e* for feminine gender.

*Adjectives: dialectal differences and historical background*

- a) In the old written monuments as well as in many dialects the “articles” that function in noun phrases with adjectival or genitival dependents are not used after the dative-genitive definite forms of nouns, cf. *njeri-ut mirë* person-GEN./DAT.SG.DEF.

good ‘good person (gen./dat.)’. In some dialects, these articles are not used after definite nominative forms, too. It remains unclear whether this feature reflects an older state or is an innovation.

- b) Tosk dialects do not have deverbal adjectives with the suffix *-shëm*, which have the meaning of the possibility or ability (*i paharrueshëm* ‘unforgettable’). These adjectives exist both in Gheg dialects and in SA.

Some Albanian adjectives are etymologically related to PIE adjectives, other adjectives are Albanian derivatives (deverbal or, less often, denominal), and there is a significant number of borrowed adjectives.

As for the origin of the grammatical layout of Albanian adjectives, a common opinion is that the prepositional article (in noun phrases with both adjectives and genitives) evolved from the PIE demonstrative pronoun and is thus etymologically identical with the postpositive definite article. Particular details of the grammaticalization of these two sets of morphemes remain a subject of a long-standing debate. Generally speaking, the origin of Albanian articulated adjectives is often compared with the development of Slavic and Baltic “long adjectives”. Unarticulated adjectives could evolve from appositive constructions of the type “noun + noun”.

Articulated adjectives have an *adverbial* function when used without article.

There is a large set of *prepositions* in Albanian; some of them are primary, and others are homonymous with adverbs (for more information on prepositions, see “Syntax”).

## Pronouns

### *Personal and demonstrative pronouns*

TABLE 12.12 DECLENSION OF THE PERSONAL AND DEMONSTRATIVE PRONOUNS

	singular				plural			
	1	2	3		1	2	3	
			m	f.			m	f
nom.	<i>unë</i>	<i>ti</i>	<i>ai</i>	<i>ajo</i>	<i>ne</i>	<i>ju</i>	<i>ata</i>	<i>ato</i>
acc.	<i>mua (më)</i>	<i>ty (të)</i>		<i>atë (e)</i>	<i>ne (na)</i>	<i>ju (ju)</i>	<i>ata (i)</i>	<i>ato (i)</i>
(gen.)/dat.	<i>mua (më)<sup>1</sup></i>	<i>ty (të)</i>	<i>atij (i)</i>	<i>asaj (i)</i>	<i>neve (na)</i>	<i>ju (ju)</i>		<i>atyre (u)</i>
abl.	<i>meje</i>	<i>teje</i>	<i>atij</i>	<i>asaj</i>	<i>nesh</i>	<i>jush</i>		<i>atyre</i>

<sup>1</sup> The (gen.)/dat. form of the personal pronoun is used very rarely in possessive constructions (only if the pronoun has its own attribute), as there are dedicated possessive pronouns that are normally used in the possessive construction (see below).

The so-called weak (clitic) forms of personal pronouns are given in parentheses (on the use of weak forms see “Syntax”).

The nominative neuter form *ata* is used very rarely; other neuter case forms don’t differ from masculine forms.

The 3rd person personal pronouns are used in the function of (or are homonymous to) distal demonstrative pronouns. Proximal demonstrative pronouns are formed by means of the formant *k(ë)* (instead of *a-*). The form of nom. m. sg. *ky* has a different vocalic pattern, while other proximal forms are distinguished from distal pronouns exclusively by the initial formant, cf. *k-jo*, *kë-të* and so on.

Dialectal differences in the domain of personal and demonstrative pronouns are rather insignificant. An older (see below) form *u* (instead *unë*) for the nom. 1 sg. pronoun is used

in many dialects as well as in the old texts. For the nom. 1 pl. and 2 pl. forms, an expansion of oblique forms (*neve, na, juve*) into the nominative slot is observed in some dialects, whereas *na* is, most likely, the original nominative form. An older form *ay* (instead *ai*) is used in the eastern part of the northern Tosk dialects. There are also some dialectal variants of the oblique case forms of the 3 sg. and 3 pl. pronouns.

### The origin

*unë* < PIE \*swom (La Piana 1949: 69, Orel 2000: 241) or < PIE \*h<sub>1</sub>eǵoH (Bopp 1854: 504–505, Matzinger 2006: 107); *në* is a kind of pronominal extension, cf. Gr. -ναί, -νή. oblique forms of the 1 sg. pronoun developed from the PIE stem \*me- (*mua* < PIE \*mēm, *meje* < PIE \*mey).

*ti* < PIE \*tuH, oblique forms reflect the same stem (*ty* < PIE \*twēm, *teje* < PIE \*tey, or under the analogical influence of *meje*).

*na* < PIE \*nos; *ne* < PIE \*noHs; *neve, nesh* (as well as *juve, jush*) could have developed under the influence from the nominal paradigm.

*ju* < PIE \*wos

The deictic elements *k(ë)*- and *a*- of the demonstrative (and 3rd person personal) pronouns must have been agglutinated to the demonstrative pronominal elements no earlier than the end of the PA period. The element *a*- could be etymologically related to the PIE deictic element \*aw-, and *k(ë)*- to the PIE element \*ko-. On the origin of demonstrative pronouns see above (“The Definite Declension”).

Clitic forms of personal pronouns evolved through the process of reduction of full forms. Contracted forms of dative and accusative pronouns should be mentioned: *ma* < *më* (1 sg. dat.) + *e* (3 sg. acc.), and so on. These forms must have emerged in the PA period, because this type of sandhi phenomena is also encountered in some nominal forms but is not a productive phonological alternation in SA.

Albanian has rather complex system of *possessive pronouns*. The first and second person possessive pronouns agree with the possessee in gender, number and case and reflect the person and number of the possessor. Most of the 1st and 2nd person possessive forms consist of a demonstrative element that is etymologically identical to Albanian article formants and a pronominal element in the first or second person, respectively. These latter elements originate from the PIE personal or possessive pronouns, although there are different opinions with respect to what kind of grammatical forms were their PIE etymons. Possessive pronouns of the 1 and 2 sg. of the possessor have analytical forms when they are used with the plural possessee. These analytical forms contain the prepositional article (see “Adjective” above) and reflect a later grammaticalization process.

The following table illustrates the declension pattern of the 1 sg. possessive pronoun:

TABLE 12.13 DECLENSION OF THE POSSESSIVE PRONOUNS

	singular of possessee		plural of possessee	
	m.	f.	m.	f.
nom.	<i>libri im</i>	<i>dora ime</i>	<i>librat e mi</i>	<i>duart e mia</i>
acc.	<i>librin tim</i>	<i>dorën time</i>	<i>librat e mi</i>	<i>duart e mia</i>
gen.-dat./abl.	<i>librit tim</i>	<i>dorës sime</i>	<i>librave të mi</i>	<i>duarve të mia</i>

*libri im* ‘my book’, *dora ime* ‘my hand’

The 3rd person possessives are analytical constructions of the type “prepositional article + demonstrative pronoun without deictic elements”; these constructions are differentiated according to the gender of the possessor: (*i tij* ‘his (masculine possessee)’, *i saj* ‘her (masculine possessee)’. These pronouns decline as adjectives.

Basic word order in the possessive phrase is “noun + possessive pronoun”.

There is a specific (probably archaic) construction that is used for the expression of the 3rd person possessive relations with some old kinship terms, inherited from PIE or borrowed from Latin, such as *atë* ‘father’, *ëmë* ‘mother’, *vëlla* ‘brother’, *motër* ‘sister’ and some others, as well as with the word *zot* ‘lord’: *i ati* ‘his, her, their father’, *e motra* ‘his, her, their sister’, *të vëllezërit* ‘his, her, their brothers’. In this construction, the element that is homonymous to the prepositive article reflects the gender and number of the possessee only. This construction alternates with the regular possessive construction. The same kinship terms also allow the prepositive use of personal pronouns of 1st and 2nd person: *im atë* ‘my father’ along with *ati im*; in this case the noun should be used in the indefinite form.

Albanian has a special *reflexive pronoun* *vetja* ‘himself’, which inflects as a noun. It marks coreference with an antecedent of any person, number and gender. The *possessive-reflexive pronoun* (*i, e*) *vet* exists only in Gheg dialects and in SA; it remains unclear whether it was ever a common-Albanian form. Etymologically, *vetja* goes back to PIE \*swe with extension \*-t- (for the morphological means of the expression of reflexive meaning, see below, “Verb”). Albanian has two reciprocal pronouns. The first, *njeri tjetrin*, lit. ‘one another.ACC.’ (and its other case forms), follows the widespread model typical of the Balkan languages. The second construction, *shoku shokun* (lit. ‘friend friend.ACC.’, *shok* < Lat. *socius*) (and its other case forms as well as phonetic variants), could have arisen under Slavic influence (cf. Old Slavic *drugъ druga*, lit. ‘another another.ACC.’, but homonymous to ‘friend friend.ACC.’). For the morphological means of the expression of the reciprocal meaning, see below, “Verb”.

Albanian has several pronouns that are derived from PIE \*k<sup>w</sup>o-/k<sup>w</sup>i- and have both *interrogative* and *relative* functions. Among these pronouns are *kush* ‘who’ < PIE \*k<sup>w</sup>os-so-, *ku* ‘where’, *kur* ‘when’, *se* ‘what’ and *që/çë* ‘what’ (for the use of the last two pronouns, see also “Syntax”). The non-inflected *që/çë* is used as a general relative pronoun; in modern SA it is being replaced by the declinable pronoun (*i, e*) *cili* (<\*t-sili) ‘which, who’, which developed from the same PIE \*k<sup>w</sup>i-.

## Numerals

### Cardinal numerals

TABLE 12.14 CARDINAL NUMERALS

1	<i>një</i> <sup>1</sup>	< PIE *sem-/sm̥- or < PIE *oyno-
2	<i>dy</i> <sup>2</sup>	< PIE *dwo-/duwo-
3	<i>tre</i> (m.), <i>tri</i> (f.)	<i>tre</i> < PIE *treyes, <i>tri</i> < *trih <sub>2</sub>
4	<i>katër</i>	< PIE *k <sup>w</sup> ₃tur-/k <sup>w</sup> ₃twor-
5	<i>pesë</i> <sup>3</sup>	<PIE *penk <sup>w</sup> e
6	<i>gjashtë</i>	<PIE *seks-t-
7	<i>shtatë</i>	<PIE *septm̥-t-
8	<i>tetë</i>	<PIE *oktoh-,t-
9	<i>nëntë</i> <sup>4</sup>	<PIE *newn-t-
10	<i>dhjetë</i>	<PIE *dekmt-

## Notes

- 1 The variant *nji* is characteristic of Gheg dialects. Cardinal numerals from ‘1’ to ‘4’ are inflected for case; *një* also has the opposition of definite and indefinite forms.
- 2 There is a special feminine form *dÿ*: in Gheg and southern Tosk dialects.
- 3 *pës* in Gheg dialects.
- 4 *nân(d)* in Gheg dialects.

The composition of cardinal numerals from ‘11’ to ‘19’ follows the model typical of Slavic languages (including Slavic languages of the Balkans), ‘x at ten’: *njëmbëdhjetë* ‘11’, *dymbëdhjetë* ‘12’ and so on.

The cardinal numerals for ‘30’ and from ‘50’ to ‘90’ are formed according to the model ‘x-ten’: *tridhjetë* ‘30’, *pesëdhjetë* ‘50’ and so on. The cardinal numerals for ‘20’ (*njëzet*) and ‘40’ (*dyzet*) are formed according to the model ‘x-twenty’, where *-zet* < PIE \*wīkṛtīh<sub>1</sub>. In most Gheg dialects *kat(ë)rdhjet(ë)* is used for ‘40’; in some Tosk dialects *trizet* is used for ‘60’, and *katërzet* for ‘80’.

The terms for ‘100’ and ‘1000’ are Latin borrowings: (*një*)*qind* < Lat. *centum*, *mijë* < Lat. *mīlia*.

### Ordinal numerals

*Ordinal numerals* are, morphologically speaking, articulated adjectives. The ordinal numerals from ‘2nd’ to ‘5th’ (as well as composite ordinal numbers that end on these digits and ‘100th’ and ‘1000th’) are formed by the means of the suffix *-të*: *i, e dytë* ‘second’, and so on. Ordinal numbers from ‘6th’ to ‘10th’ (and composite ordinal numbers that end on these digits) are differentiated from the corresponding cardinal numbers only by the presence of the prepositive article: *i, e gjashtë* ‘sixth’ (*-të* is here a result of the simplification of the sequence “*-të* of the stem + suffix *-të*”). Ordinal number *i, e parë* ‘first’ < PIE \*pr̥H-wo-.

### Verb system

#### General remarks

The Albanian verb system is very complex with respect to both the number of categories and the number of conjugational types and subtypes. Grammatical meanings are expressed both synthetically and analytically. In the synthetic forms, vocalic and consonantal alternations play a significant role along with affixation. Analytical constructions are formed by means of conjugated auxiliary verbs and clitic particles.

The verb has the following grammatical categories: person, number, tense, aspect (partly overlapping with tense), mood, evidentiality (admirative) and voice.

The person-number (1st, 2nd and 3rd person, singular and plural) distinctions are expressed by the verb endings and, in some conjugational types, by vocalic and consonantal alternations. Different tense-aspect-mood forms and sometimes different conjugational types often have different sets of person-number endings.

The organization of TAM categories is very complex, because all the moods (except imperative) have temporal distinctions. Some forms that are traditionally described as “tenses” are in fact opposed to each other in terms of their aspectual meaning (first of all,

the *aorist* and *imperfect*). Albanian has *indicative, subjunctive, conditional, optative* and *imperative* moods.

*Indicative* has the richest temporal system. It includes *three synthetic tenses*, which form the bulk of the Albanian temporal system: *present, aorist* and *imperfect*. Present and aorist are differentiated by their stems (not for all verbs) and paradigms of endings. The forms of the imperfect are derived from the present stem (or from one of its variants in cases of *morphophonemic* alternations) and have a special set of endings; historically, imperfect endings in the active voice are similar to aorist endings. Semantically, the aorist is close to a prototypical perfective past, and the imperfect to a prototypical imperfective past as they feature in many Indo-European languages. The *perfect subsystem* has forms of the *perfect* and two *pluperfects*. In the active voice these are expressed by combinations of the auxiliary verb *kam* 'to have' in the corresponding tense form (present for the perfect, imperfect and aorist for the two pluperfects) and the participle of the main verb. The perfect has a bundle of meanings characteristic of this aspectual form from the typological point of view, such as perfect of the current relevance, experiential perfect and so on. The two pluperfects occupy the semantic niche that is typical of the traditional pluperfects: relative tense precedence or absolute temporal distance. The difference between the two pluperfects has not been studied in sufficient detail, but the aorist pluperfect is clearly a rather rare form. Albanian has also a set of so-called *supercompound* forms, which are characterized by the use of the *compound (perfect or pluperfect)* forms of the auxiliary verbs. These supercompound forms have restricted dialectal distributions but are represented, although very poorly, in standard texts. Semantically, they express various shades of temporal remoteness. The *future subsystem* is formed in SA by the means of a special future particle *do* and the subjunctive of the main verb. This subsystem includes *future* ("do + present subjunctive"), *future in the past* ("do + imperfect subjunctive"), *future perfect* ("do + perfect subjunctive") and *future perfect in the past* ("do + pluperfect subjunctive"). Albanian also has a "parallel" future subsystem with an additional necessitive semantic shade; these forms are formed by means of the verb 'to have' (in the present, imperfect and aorist) and the so-called new infinitive (see below).

SA also has two additional analytical constructions with *aspectual* meanings. The first construction consists of *jam* 'to be' (present, imperfect or aorist) + analytical gerund (see below) and has a progressive meaning: *jam duke punuar* be.1SG.PRES. GERUND work.PTCP. 'I am working'. The second construction consists of the particle *po* and the present or imperfect form of the verb; it expresses the progressive meaning and relates the time of action to a certain reference time, either the moment of speech or a certain interval of time in the past: *po punoj* PROG. work.1SG.PRES. 'I am working (right now)'.

The *subjunctive* mood is formed with the special particle *të* (which is derived from an old conjunction?) and synthetic or analytical verbal forms that partly coincide with those of the indicative mood; for the absolute majority of verbs only the 2 sg. and 3 sg. subjunctive present tense forms of the active voice are different from the corresponding forms of the indicative mood. The subjunctive is used in complement clauses as an analog of the lacking infinitive (see below), in the main clauses in imperative and optative functions; it is also used in various hypothetical constructions (see Breu 2010: 465–466). The subjunctive has *present, imperfect, perfect* and (imperfect) *pluperfect* forms. In the old texts the subjunctive is used without *të* after the negation particle *mos* (often) and after the possibility particle *mundë, munë* (nearly always).

The *conditional* is formed by combining the subjunctive with the prepositional particle *do*. The main context of the use of the conditional is the apodosis of hypothetical periods

(see Breu 2010: 467). The conditional has present (“do + imperfect subjunctive”, homonymous to the future in the past) and past (“do + pluperfect subjunctive”, homonymous to the future perfect in the past) forms; thus, the paradigm of forms with *do* can be viewed as a single *future-conditional* system.

Additionally, there are some analytical constructions formed with various particles and subjunctive forms. They have various modal meanings (jussive, possibilitive and so on), but their status as real grammatical moods is questionable.

The *optative* mood has present and perfect forms. The present optative is a synthetic form that has a special stem, which in most classes of verbs is derived from the aorist stem, and a special set of endings. The perfect optative is an analytical form, consisting of the optative of the auxiliary and a participle of the main verb. In SA, the optative is mostly used in various fossilized or semi-fossilized expressions (blesses, curses and so on); in the optative function proper, which is restricted and stylistically marked; and in the protasis of hypothetical periods.

The *imperative* mood has two synthetic forms, the 2 sg. and the 2 pl. form.

The basic meaning of the SA *admirative* is mirative, but it also has a secondary commentative function. Admirative distinguishes at least two moods, *indicative* and *subjunctive* (the forms of the *conditional admirative* are very rare), and some tense forms. The synthetic forms of the *present* and *imperfect* indicative admirative emerged through amalgamation of inverted perfect and pluperfect forms, respectively, cf. *bëkam* [do.ADM. PRES. 1SG.] < *bë(rë) kam* [do.PART. have.PRES. 1SG.]. From a synchronic point of view they have a special stem and a special set of endings. The perfect and pluperfect admirative forms are formed by the admirative forms of the auxiliary verb and the participle of the main verb. The *subjunctive-admirative* forms (present, imperfect and pluperfect) are formed with the particle *të* and corresponding admirative verbal forms.

There are two *voices* in Albanian, *active* and *non-active*. Albanian non-active forms can have different semantic interpretations: passive, reflexive, reciprocal, anti-causative etc. Passives are regular (inflectional) voice forms of transitive verbs, whereas other functions are valency-reducing derivations of active verbs; some of these derived verbs significantly deviate from their active counterparts in terms of their meaning. There is also a rather small number of semantically active deponent verbs in Albanian.

The formal structure of the Albanian non-active voice system is extraordinarily complex. There are three morphological devices that can signal non-activeness. At least in SA, these three devices are rigidly distributed among the tense-modal forms of verbs.

- a) A set of special endings; this device is employed for the present, imperfect and future indicative as well as for the present and imperfect subjunctive and present of the conditional (= future in the past indicative).
- b) A construction that consists of the particle *u* and an active verb form; this device is employed for the aorist indicative, present and imperfect admirative, present optative and imperative; besides, for most Albanian verbs the 3 sg. aorist form in this construction is formally distinct from the corresponding active form.
- c) A corresponding finite form of the verb *jam* ‘be’ + participle; this device is employed for the forms of the perfect system.

In modern standard Albanian, these three types of non-active forms can express all of the meanings that belong to the non-active semantic spectrum as outlined above: reflexive, passive, anti-causative and so on.



### *Non-finite forms*

Albanian has a single *participle* formed from both transitive and intransitive verbs. The participle is used in the analytical verbal constructions of the “perfect system” and in other analytical constructions (see below), as a predicative attribute and in relative constructions. On the adjectivization of participles, see above (“Adjective”). The (non-adjectivized) participle is a non-inflected form. Besides, there are some analytical non-finite constructions that are formed by means of a participle. The most widespread among these constructions are the following:

- *duke* + participle (“gerund”). This construction chiefly expresses the meaning of simultaneity. It has a non-active form (*duke* + *u* + participle) and a rare perfect active (*duke* + *pasë* + participle) and passive (*duke* + *qenë* + participle) forms.
- *pa* + participle (“privative”). This construction behaves syntactically just like the gerund and has the same voice and temporal forms as the gerund.
- *për* + *të* + participle (“new infinitive”). This construction is mainly used with the final (= purpose) meaning; it has a non-active counterpart.
- *më* + *të* + participle (“absolutive”). This construction’s main meaning is temporal precedence. The new infinitive and the absolutive are in fact formed with the help of the deverbal noun, which is in turn derived from the participle.
- The ablative of the deverbal nouns is used in a function that is close to that of the infinitive after the phasal verbs. This form is mostly used after verbs that denote termination or cancellation of the action: *mbaroi së fol-ur-i* finished ART.ABL. speak-PART.-ABL. ‘s/he stopped speaking’.

### *Conjugational classes*

There is no uniform and commonly accepted system for the classification of the numerous SA conjugational classes. The majority of the classification attempts (see, first of all, Boretzky 1977, Buchholtz & Fiedler 1987) take as a basis for the first-level classification the form of the present stem and, accordingly, the characteristics of the conjugation in the present (for a different approach, see Schumacher & Matzinger 2013).

Under this approach, three major classes may be distinguished: (a) the verbs with the stem ending in a vowel that take endings in the singular present indicative, (b) the verbs with the stem ending in a consonant, and (c) the verbs with the stem ending in a vowel that do not take endings in the singular present indicative (this class is smallish). Besides these classes, there are some verbs that can be synchronically viewed as exceptions.

Further classification of the conjugational subclasses is based both on the peculiarities of the conjugation in present (first of all on the consonantal and vocalic alternations) and on the relation between the present and aorist stems; sometimes other verbal forms – imperative, participle and so on – are also taken into account. The present and the aorist stem types are two subsystems that are connected to each other by many-to-many relations. As a consequence, an independent classification based on the form of the aoristic rather than the present stem can also be elaborated (see Boretzky 1977). The vocalic and consonantal alternations are also independent of the verb class as established by the stem formation processes, although some kinds of alternations as well as some types of the aorist stems are not encountered in certain classes as defined in terms of present stem formation.

The only productive conjugational subtype of modern SA is the 1st class with the stem ending in *-o* (type *punoj*).

Table 12.15 summarizes examples of major conjugational classes.

**TABLE 12.15 VERB CONJUGATION (THE MAJOR CONJUGATIONAL CLASSES)**

		I. <i>punoj</i> 'to work'	II. <i>kap</i> <sup>1</sup> 'to seize'	III. <i>pi</i> 'to drink'
Active				
Indicative present	1 sg.	<i>punoj</i>	<i>kap</i>	<i>pi</i>
	2 sg.	<i>punon</i>	<i>kap</i>	<i>pi</i>
	3 sg.	<i>punon</i>	<i>kap</i>	<i>pi</i>
	1 pl.	<i>punojmë</i>	<i>kapim</i>	<i>pimë</i>
	2 pl.	<i>punoni</i>	<i>kapni</i>	<i>pini</i>
	3 pl.	<i>punojnë</i>	<i>kapin</i>	<i>pinë</i>
Indicative imperfect	1 sg.	<i>punoja</i>	<i>kapja</i>	<i>pija</i>
	2 sg.	<i>punoje</i>	<i>kapje</i>	<i>pije</i>
	3 sg.	<i>punonte</i>	<i>kapte</i>	<i>pinte</i>
	1 pl.	<i>punonim</i>	<i>kapnim</i>	<i>pinim</i>
	2 pl.	<i>punonit</i>	<i>kapnit</i>	<i>pinit</i>
	3 pl.	<i>punonin</i>	<i>kapnin</i>	<i>pinin</i>
Indicative aorist	1 sg.	<i>punova</i>	<i>kapa</i>	<i>piva</i>
	2 sg.	<i>punove</i>	<i>kapë</i>	<i>pive</i>
	3 sg. <sup>2</sup>	<i>punoi</i>	<i>kapi</i>	<i>piu</i>
	1 pl.	<i>punuam</i>	<i>kapëm</i>	<i>pimë</i>
	2 pl.	<i>punuat</i>	<i>kapët</i>	<i>pitë</i>
	3 pl.	<i>punuan</i>	<i>kapën</i>	<i>pinë</i>
Subjunctive Present <sup>3</sup>	2 sg.	<i>të punosh</i>	<i>të kapësh</i>	<i>të pish</i>
	3 sg.	<i>të punojë</i>	<i>të kapë</i>	<i>të pijë</i>
Optative present	1 sg.	<i>punofsha</i>	<i>kapsha</i>	<i>pifsha</i>
	2 sg.	<i>punofsh</i>	<i>kapsh</i>	<i>pifsh</i>
	3 sg.	<i>punofië</i>	<i>kapfië</i>	<i>piffië</i>
	1 pl.	<i>punofshim</i>	<i>kapshim</i>	<i>pifshim</i>
	2 pl.	<i>punofshit</i>	<i>kapshit</i>	<i>pifshit</i>
	3 pl.	<i>punofshin</i>	<i>kapshin</i>	<i>pifshin</i>
Admirative present	1 sg.	<i>punuakam</i> <sup>4</sup>	<i>kapkam</i>	<i>pikam</i>
	2 sg.	<i>punuake</i>	<i>kapke</i>	<i>pike</i>
	3 sg.	<i>punuaka</i>	<i>kapka</i>	<i>pika</i>
	1 pl.	<i>punuakemi</i>	<i>kapkemi</i>	<i>pikemi</i>
	2 pl.	<i>punuakeni</i>	<i>kapkeni</i>	<i>pikeni</i>
	3 pl.	<i>punuakan</i>	<i>kapkan</i>	<i>pikan</i>
Admirative imperfect	1 sg.	<i>punuakësh</i> <sup>4</sup>	<i>kapkësh</i>	<i>pikësh</i>
	2 sg.	<i>punuakeshe</i>	<i>kapkësh</i>	<i>pikësh</i>
	3 sg.	<i>punuakësh</i>	<i>kapkësh</i>	<i>pikësh</i>
	1 pl.	<i>punuakëshim</i>	<i>kapkëshim</i>	<i>pikëshim</i>
	2 pl.	<i>punuakëshit</i>	<i>kapkëshit</i>	<i>pikëshit</i>
	3 pl.	<i>punuakëshin</i>	<i>kapkëshin</i>	<i>pikëshin</i>
Imperative	2 sg.	<i>puno</i>	<i>kap</i>	<i>pi</i>
	2 pl.	<i>punoni</i>	<i>kapni</i>	<i>pini</i>
Non-active				
Indicative present	1 sg.	<i>punohem</i>	<i>kapem</i>	<i>pihem</i>
	2 sg.	<i>punohesh</i>	<i>kapesh</i>	<i>pihesh</i>
	3 sg.	<i>punohet</i>	<i>kapet</i>	<i>pihet</i>
	1 pl.	<i>punohemi</i>	<i>kapemi</i>	<i>pihemi</i>
	2 pl.	<i>punoheni</i>	<i>kapeni</i>	<i>piheni</i>
	3 pl.	<i>punohen</i>	<i>kapen</i>	<i>pihen</i>

Indicative imperfect	1 sg.	<i>punohesha</i>	<i>kapesha</i>	<i>pihesha</i>
	2 sg.	<i>punoheshe</i>	<i>kapeshe</i>	<i>piheshe</i>
	3 sg.	<i>punohej</i>	<i>kapej</i>	<i>pihej</i>
	1 pl.	<i>punoheshim</i>	<i>kapeshim</i>	<i>piheshim</i>
	2 pl.	<i>punoheshit</i>	<i>kapeshit</i>	<i>piheshit</i>
	3 pl.	<i>punoheshin</i>	<i>kapeshin</i>	<i>piheshin</i>
Indicative aorist <sup>5</sup>	3 sg.	<i>u punua</i>	<i>u kap</i>	<i>u pi</i>
Participle		<i>punuar</i>	<i>Kapur</i>	<i>pirë</i>

In accordance with the Albanian lexicographic tradition, I use the 1 sg. indicative present form as the citation form.

- 1 In SA, there are three verbs with the stem ending in a consonant, namely *eci* 'to go', *iki* 'to leave, to depart' and *hipi* 'to raise', which take a special set of endings in the singular present indicative: 1 sg. *-i*, 2–3 sg. *-ën*.
- 2 The ending *-i* is used after *-o-* and consonants (except *-k-*, *-g-*, *-h-*); the ending *-u* is used elsewhere.
- 3 Only those forms that deviate from the indicative pattern are given.
- 4 The variant forms: *punokam*, *punokësha* etc.
- 5 Only those forms that deviate from the active forms are given.

### Major types of the aorist stems

- The aorist stem is identical to the present stem. The verbs of this type belong to all three conjugational classes. If the stem ends in a vowel the consonant *-v-* is inserted in the 1 sg. and 2 sg. forms.
- The aorist stem is characterized by the extension *-(j)t-* in both singular and plural, cf. 1 sg. pres. *dī* – 1 sg. aor. *dīta* 'to know'. This type of correspondence is characteristic of some verbs from the 1st and 3rd classes.
- The aorist stem is characterized by the extension *-r-* in the singular, cf. 1 sg. pres. *běj* – 1 sg. aor. *běra* 'to do'. The correspondence of this type is encountered in some verbs from the 1st and 3rd classes. Historically *-r-* (Gheg *-n-*) was also present in the present stem.
- The aorist stem ends in a vowel, whereas the present stem has an extension *-s-/-t-*, cf. 1 sg. pres. *vrās* – 1 sg. aor. *vrava* 'to kill'. This type of correspondence is encountered in some verbs of the 2nd class.

### Major vocalic and consonantal alternations

Below I discuss only those alternations that are realized in a considerable number of verbs.

- a/-e-*: *e* is used in the 2 sg. and 3 sg. present indicative active forms of some verbs of the 2nd class.
- e- (-je-, -ie-)/-i-*: *i* is used in the 2 pl. present indicative (and subjunctive, future and so on) active, in the imperfect active, in non-active synthetic forms, and in the imperative. This alternation is characteristic of verbs of the 2nd and 3rd classes (and for one verb of the 1st class).
- a-*, *-e-/-o-*: *-a-*, *-e-* (*e* further alternates with *-i-*; see above) is encountered in all forms except the aorist; it is typical of some verbs of the 2nd class.
- o-/-ua-*; *-e-/-ye-*: the diphthongs are used in the plural forms of the aorist (and in the 3 sg. of the aorist non-active), in the admirative and in the participle. This alternation

- is characteristic of the absolute majority of verbs of the 1st class that end in *-o* and of the majority of verbs of the 1st class that end in *-e*.
- ua/-o-*; -*ye/-e-*: *o* resp. *e* is encountered the singular forms of the aorist (except 3 sg. of the aorist non-active) and in the optative of those verbs of the 1st class that have a present stem ending in *-ua, ye*. In the synthetic forms of the non-active these verbs demonstrate *u* resp. *e*. This alternation pattern is found in a small group of the verbs of the 1st class.
  - t/-s-*: the alternant *s* is used in the 1 sg. present indicative active and in the 2 sg. and 3 sg. present subjunctive active and, often parallel to the forms with *-t-*, in the 2 pl. present indicative active, in the imperfect indicative active and in the 2 pl. imperative. The alternation is quite typical of the verbs of the 2nd class.
  - k/-q-*; -*g-* / -*gj*: palatal consonants are used in the 2 pl. present indicative (and subjunctive, future and so on) active, in the imperfect active, in non-active synthetic forms, in the imperative and in the aorist. This alternation pattern is typical of some verbs of the 2nd class.

The conjugation of the two most important irregular verbs is illustrated in Table 12.16.

TABLE 12.16 CONJUGATION OF THE VERBS *JAM* 'TO BE' AND *KAM* 'TO HAVE'

		<i>jam</i> 'to be'	<i>kam</i> 'to have'
Indicative present	1 sg.	<i>jam</i>	<i>kam</i>
	2 sg.	<i>je</i>	<i>ke</i>
	3 sg.	<i>është</i>	<i>ka</i>
	1 pl.	<i>jemi</i>	<i>kemi</i>
	2 pl.	<i>jeni</i>	<i>keni</i>
	3 pl.	<i>janë</i>	<i>kanë</i>
Indicative imperfect	1 sg.	<i>isha</i>	<i>kisha</i>
	2 sg.	<i>ishe</i>	<i>kishe</i>
	3 sg.	<i>ish(te)</i>	<i>kish(te)</i>
	1 pl.	<i>ishim</i>	<i>kishim</i>
	2 pl.	<i>ishit</i>	<i>kishit</i>
	3 pl.	<i>ishin</i>	<i>kishin</i>
Indicative aorist	1 sg.	<i>qeshë</i>	<i>pata</i>
	2 sg.	<i>qe</i>	<i>pate</i>
	3 sg. <sup>2</sup>	<i>qe</i>	<i>pati</i>
	1 pl.	<i>qemë</i>	<i>patëm</i>
	2 pl.	<i>qetë</i>	<i>patët</i>
	3 pl.	<i>qenë</i>	<i>patën</i>
Subjunctive present	1 sg.	<i>të jem</i>	<i>të kem</i>
	2 sg.	<i>të jesh</i>	<i>të kesh</i>
	3 sg.	<i>të jetë</i>	<i>të ketë</i>
	3 pl.	<i>të jenë</i>	<i>të kenë</i>
Optative present	1 sg.	<i>qofsha</i>	<i>paça</i>
	2 sg.	<i>qofsh</i>	<i>paç</i>
	3 sg.	<i>qoftë</i>	<i>pastë</i>
Admirative present	1 sg.	<i>qenkam</i>	<i>paskam</i>
Participle		<i>qenë</i>	<i>pasur</i>

Among other irregular verbs are the following (I provide only the 1 sg. present indicative form, the 1 sg. aorist indicative form and the participle): *bie – rashë – rënë* 'to fall'; *bie – prurë* 'to bring'; *dua – desha – dashur* 'to want, to love'; *ha – hëngra – ngrënë*

‘to eat’; *jap – dhashë – dhënë* ‘to give’; *lë – lashë – lënë* ‘to leave’; *rri – ndenja – ndenjur* ‘to sit, to stay’; *shoh – pashë – pare* ‘to see’; *them – thashë – thënë* ‘to say’; *vdes – vdiqa – vdekur* ‘to die’; *vete – vajta – vajtur* ‘to go’; *vij – erdha – ardhur* ‘to come’.

### *Basic patterns of dialectal variation*

- 1) Admirative forms are almost absent in the Arbresh and Arvanitika dialects, as well as in the southern Tosk Cham dialects.
- 2) The supercompound forms are alien to the northern Tosk dialects.
- 3) In old Albanian written texts, as well as in a portion of modern Gheg dialects (mainly in the dialects of the eastern northern group), several intransitive verbs are used in the perfect tenses with the auxiliary ‘to be’, cf. *jam shkue* (I.am go.PART.) ‘I have walked,’ instead of SA *kam shkuar* (I.have go.PART.).
- 4) In Gheg dialects there is a special form of the analytical infinitive that is formed according to the model *me* + participle (*me shkue* ‘to go’). Most likely, the form is a Gheg innovation.
- 5) In Gheg dialects there is a special form of the future, which is formed according to the model ‘have + analytical infinitive’, cf. *kam me shkue* have.1SG.PRES. INF. PART. ‘I shall go’; this form coexists with the SA form of the future, which apparently has radiated from the Tosk area. Other Gheg forms of the future and conditional subsystems can also be formed according to this model (see Breu 2013: 43). In Arbresh dialects there is a “hybrid” future form (*kam* + subjunctive).
- 6) There are various dialectal forms of the gerund participle (for example, *tue* in Gheg).
- 7) There are numerous dialectal differences in the field of participle formation. Derivational types are more variegated in Gheg dialects than in Tosk dialects (see below). Besides, in Gheg dialects there is a special short form of the participle that is used in analytical constructions.
- 8) In some Gheg dialects (especially in the eastern part of the northern and central Gheg dialects) there is a tendency towards non-differentiation of the 2 sg. and 3 sg. present indicative active forms with corresponding subjunctive forms.
- 9) In most Gheg dialects (except South Gheg dialects) the imperfect is formed by means of the suffix *-sh-* (Gheg *punojsha* ‘I worked’). The almost total absence of such forms in the old texts is indicative of the recent origin of this innovation. There are also many other dialectal peculiarities in the imperfect endings (for example, 3 sg. *-ke* in eastern northern Gheg dialects).
- 10) The imperfect passive is formed in the majority of Gheg dialects by means of the particle *u* and not synthetically as it is in SA (Gheg *u lajsha* REFL. wash.1SG.IMP. ACT. vs. SA and Tosk *lahesha* wash.1SG.IMP.PASS. ‘I was washed’). In a few dialects (Gheg, southern Tosk and Arbresh) as well as in the old texts a parallel non-active perfect construction is also encountered, which is formed with the particle *u*: “*u* + perfect (pluperfect, aorist 2) active”.
- 11) It is often the case that many verbs shift in various dialects from one conjugational class to another. For example, in some northern Tosk dialects there is a tendency towards the extension of the aorist stems in *-t-* at the expense of those verbs that have aorist stems with *-v-* in SA. In the old texts, the conjugation type that is typical of verbs *ik*, *ec* and *hip* in SA is extended to a significant number of verbs of the 2nd conjugational class (see de Vaan 2014); thus, these verbs constitute a special

subclass of the 2nd class in the old texts. In some Gheg dialects, the verbs of the 3rd class conjugate in the present according to the model of the 1st class.

- 12) There are also dialectal differences in conjugation that are due to phonetic processes. For example, the SA and Tosk *-r-* in the forms of the aorist corresponds to *-n-* in Gheg; cf. SA and Tosk 1 sg. aor. act. *bëra* ‘I did’ – Gheg *bāna*. In a part of the Tosk dialectal area the present and imperfect endings with *-j(-)* retain the archaic form *-nj(-)*, cf. *punoj* ‘I work’ as opposed to the SA *punoj*.
- 13) There are some archaic endings in the old texts. They are important for the reconstruction of the history of the Albanian conjugation; see below on some of them.

### *Verb system: the origin*

#### Present stems

The PIE verb system has been significantly reshaped in Albanian. Major conjugational types of SA often encompass verbs that evolved from verbs that have different types of PIE stems. Only three verbs that use *-m* as their 1 sg. present indicative and subjunctive ending can be treated as the continuation of the PIE athematic type: SA *jam* ‘to be’ (< PIE \*h<sub>1</sub>es-/h<sub>1</sub>s-), SA *them* ‘to say’ (< PIE \*keHs-) and SA *kam* ‘to have’ (< PIE \*keh<sub>2</sub>p-); the latter verb could have been influenced by the verb *jam*. All other (inherited) verbs evolved from PIE thematic types, with or without suffixes.

- “Pure” thematic type; verbs that originated from this type mostly belong to the SA second conjugational class: SA *bređh* ‘to roam’ < PIE \*b<sup>h</sup>red<sup>h</sup>-e/o-; *vjedh* ‘to steal’ < PIE \*weg<sup>h</sup>-e/o-.
- Suffix *\*-ske/-o-*: SA *njoh* ‘to know, acquaint’ < PIE \*ǵnh<sub>3</sub>-ske-/o- (PIE \*sk > SA *-h*, see “Phonology”).
- Suffix/infix *\*-ne/o-*: SA *dal* ‘go out’ < \*d<sup>h</sup>l-ne/n-h<sub>1</sub>-; SA *bind* ‘to convince’ < PIE \*b<sup>h</sup>i-ne/n-d<sup>h</sup>-.
- The present suffix *\*-ye/o-* and causative-iterative suffix *\*-eye/o-* play a key role in the development of the Albanian conjugational system. Some verbs of the SA 2nd and 3rd conjugational classes are formed with these suffixes: SA *bie* ‘to bring’ < PIE \*b<sup>h</sup>er-ye-; SA *ndez* ‘to burn’ < PIE \*dog<sup>wh</sup>-eye/o-; SA *vesh* ‘to put on’ < PIE \*wos-eye/o-. When the verb stem ended in *-t-*, attachment of this suffix resulted in the alternation *-s/-t-*. For example SA 1 sg. *flas* ‘to speak’ is a denominal formation based on the Latin borrowing *fabula* > SA *fjalë* < \*-at-yō, 2 sg. *flet* < \*-at-is < \*at-yes (see Matzinger 2006: 122). This alternation became a productive pattern in a certain period of Albanian language history (it is typical, for example, of verbs borrowed from Slavic) and is now found in a large number of verbs of the 2nd class. The agglutination of the suffix to the nasal presents (for example, SA *bëj* ‘to do’ < PIE \*b<sup>h</sup>-ŋ-h<sub>2</sub>-ye/o-) resulted in the development of the suffix *-nyo/e-*, which has become the main tool for verb derivation and thus eventually formed the 1st conjugational class. Many denominal verbs were formed with this suffix (SA *punoj* ‘to work’ from SA *punë* ‘work’ < PIE \*pudneh<sub>2</sub>). In addition, some verbs from other conjugational classes demonstrate or demonstrated a tendency to shift into this class (see, for example, SA *qaj* ‘to cry’, which belonged to the 3rd class in Buzuk’s book). Starting from the early Albano-Latin contacts, most borrowed verbs are adapted as verbs of the 1st class (ending in *-oj* or, more rarely, in *-ej*).

*Present endings*

## Active indicative

- 1 sg.: athematic SA *-m* < PIE *\*-mi*; thematic SA *-Ø* < PIE *\*-ō* (*\*-oh<sub>2</sub>*); SA *-j*, OA *-nj* (in the 1st class) < PIE *\*-n-yō*; SA *-s* (in the 2nd class) < PIE *\*-t-yō*.
- 2 sg. and 3 sg.: The zero endings of the 2 sg. and 3 sg. of the thematic conjugation can reflect either the PIE injunctive (“secondary”) endings *\*-es*, *\*-et*, or the “primary” endings *\*-esi*, *\*-eti*. Umlaut in some Albanian verbs shows that the immediate stage of development was *\*-is*, *\*-it*. The forms ending in *-n* (1st class) and *-t* (2nd class) indicate at the loss of the *-y-* in the suffix (see Schumacher & Matzinger 2013: 253). *-të* in the 3 sg. of *-mi* verbs must have developed from PIE *\*-ti*.
- 1 pl.: thematic SA *-më* < PIE *\*-omes*; athematic SA *-mi* < PIE *\*-mes*.
- 2 pl.: SA *-ni* (in the old texts and in dialects a more archaic ending *-i* is used). The ending *-ni* is a later analogical development influenced by the verbs of the 1st class; *-i* could have evolved due to the influence from the non-active paradigm (Klingenschmitt 1994: 226).
- 3 pl.: SA *-në* < PIE *\*-(o)nti* (an alternative interpretation is given in Schumacher & Matzinger 2013: 51–52, 215).

## Non-active indicative

Non-active present endings continue PIE medial present endings. If the verb’s stem ends in a consonant (2nd class), these endings are attached directly to the stem; if the verb’s stem ends in a vowel (3rd class), a connecting consonant *-h-* is inserted (it is either a later anti-hiatus consonant or a reflex of PIE *\*-sk-*). The same *-h-* is also found in the verbs of the 1st class, but in the old texts and in some dialects non-active endings are attached to a stem ending in *-n*: SA *bëhet* ‘s/he becomes (3 sg. pres. non-act.)’ vs. OA *bënet*. Non-active forms, both present and imperfect, of the verbs with vocalic alternations have the stem variant with *-i-* (as well as with final palatal consonants *-q*, *-gj*). Some peculiarities of the spelling of non-active present forms in the old Gheg texts make it possible to conclude that in OA the stress fell on the thematic vowel (see Schumacher & Matzinger 2013: 127–131).

- 1 sg. SA *-em* < PIE *\*-o-may*.
- 2 sg. SA *-esh*; OA *-ë* < PIE *\*-e-soy*.
- 3 sg. SA *-et*, OA *-etë* < PIE *\*-e-toy*.
- 1 pl. SA *-emi* < PIE *\*-o-med<sup>h</sup>h<sub>2</sub>* (?).
- 2 pl. SA *-eni*, OA *-i* < PIE *\*-e-d<sup>h</sup>we* (?).
- 3 pl. SA *-en*, OA *-enë* < PIE *\*-o-ntoy*.

The Albanian *present active subjunctive*, at least those forms that are different from indicative forms, resulted from the development of the PIE subjunctive.

- 2 sg. SA *-sh* < PIE *\*-e-esi*; 3 sg. SA *-ë* < PIE *\*-e-et(i)*; PIE *\*-e-si*, *\*-e-ti* for athematic verbs.

## Imperfect active

Imperfect forms are based on the present form. The verbs with vocalic alternations have in the imperfect the stem variant with *-i-* (as well as with final palatal consonants *-q*, *-gj*).

The imperfect demonstrates a very high level of diatopic and diachronic variability; in old Albanian various forms of the imperfect are encountered in texts by different writers. In what follows I will rely on the reconstruction of late PA imperfect endings offered by St. Schumacher (Schumacher & Matzinger 2013: 134–140). It is generally accepted that imperfect endings in Albanian developed from the PIE “secondary” endings.

- 1 sg. SA *-ja*, PA *\*-/ə/* < PIE *\*-(o)-m/-m̃*. In SA (and, partly, in old texts) the modern endings with the element *-j-* (dialectal *-nj-*) in 1 sg. and 2 sg. and with *-n-* in the plural have extended under the influence of the verbs of the 1st class verbs and also of the OA verbs of the *ikinj* type, see above; *-a* in the ending *-ja* could have arisen under the influence of the corresponding aorist form, see below.
- 2 sg. SA *-(j)e*, PA *\*-e* < PIE *\*-(e)-s*.
- 3 sg. SA *-te*, PA *\*-Ø*, *\*-j* < PIE *\*-(e)-t*; SA *-te* is a relatively late extension.
- 1 pl. SA *-nim*, PA *\*-im* < PIE *\*-(o)-mo/e*.
- 2 pl. SA *-nit*, PA *\*-itə* < PIE *\*-(e)-te*.
- 3 pl. SA *-nin*, PA *\*-(i)nnə* < PIE *\*-(o)-nti* (an alternative explanation is suggested in Schumacher & Matzinger 2013: 215)

### Imperfect non-active

The endings of the imperfect non-active did not evolve continuously from the PIE for-mants; rather, the paradigm has been deeply reshaped according to the model of the conjugation of the verb *jam* ‘to be’ in the imperfect active. This pattern thus shows an interesting parallel to the pattern in Modern Greek (Schumacher & Matzinger 2013: 143, Fiedler 1981). The accent patterns of the imperfect non-active in old Gheg texts resemble those of the present non-active.

### Aorist stems

The Albanian aorist represents a highly heterogeneous phenomenon. From a historical point of view the following types can be distinguished:

- The verbs that use *-shë* in the 1 sg. aor. act.: *qeshë* ‘I was’, *thashë* ‘I spoke’, *lashë* ‘I left’, *dhashë* ‘I gave’, see above the list of the irregular verbs (pp. 589–590). These forms have been traditionally viewed as directly descending from the PIE sigmatic aorist. According to Klingenschmitt 1994 (see also Matzinger 2006: 128–129), they are innovations that developed under the analogical influence of the imperfect of athematic verbs (Buzuk’s *jeshë* ‘I was (impf.)’ < PIE *\*h<sub>1</sub>es-m̃*).
- By contrast, the aorists in *-v-* (1st and 3rd classes) must have developed from the PIE sigmatic type, in which the intervocalic *\*-s-* has dropped and an anti-hiatus *-v-* has emerged (Klingenschmitt 1994, Matzinger 2006: 129, Schumacher & Matzinger 2013: 148–154). Another hypothesis with respect to the origin of the *-v-* aorist assumes that it is a form that is parallel to Latin *-vī-* perfects (Orel 2000: 199).
- There is no commonly accepted view on the etymology of aorists in *-t-*. A possible suggestion is that it evolved from the periphrastic construction consisting of the verbal adjective (with *\*-to-* suffix) and *esse* (Schumacher & Matzinger 2013: 154–158, Klingenschmitt 1994: 231).



- Among the Albanian root aorists that are especially interesting are some 30 verbs that have an alternation between *-a-* or *-e-* in the present and *-o-* (< PIE \*ē) in the aorist: SA *hedh* ‘I throw’ : *hodha* ‘I threw’. It is usually assumed that these aoristic forms have evolved from the perfect with a lengthened root vowel; this pattern is then similar to analogical patterns in Latin, German, Celtic and Tocharian (for a detailed analysis, see Schumacher & Matzinger 2013: 161–173; for an interpretation of these aorists as “Narten-imperfects,” see Jasanoff 2012).

#### The aoristic endings in SA

- 1 sg. *-a*, 2 sg. *-e*, 3 sg. *-i/-u*, 1 pl. *-m(ě)*, 2 pl. *-t(ě)*, 3 pl. *-n(ě)*, have evolved from the PIE secondary endings (similarly to imperfect active endings). The 1 sg. and 3 sg. call for some comments.
  - 1 sg. SA *-a* is probably the result of contraction of *-ě-* and *-e* (on parallel contractions in nominal and pronominal system, see above). *-ě-* could have developed from PIE \**-m̥* or from the perfect ending \**-h<sub>2</sub>a* (see Matzinger 2006: 129). *-e* may be a reflex of a deictic element.
  - 3 sg. *-i/-u* (on the distribution of these variants see above) goes back to a deictic element that is identical to the definite article (and to the pronominal element of the demonstrative pronoun). This element has entered the paradigm relatively recently (in the OA period).

#### Aorist non-active

Non-active particle *u* < PIE \**swe*.

- 1 sg. *-shě*, attested in old texts and in some dialects, has evolved under the influence of the imperfect paradigm.
- 3 sg. SA *-Ø* – the active ending without the deictic particle.

All other non-active forms don’t differ from the active forms.

#### Optative

The optative is traditionally considered an innovation formed on the basis of aorist stems (see e.g. Matzinger 2006: 132; there is also a proposal that the Albanian optative evolved from the PIE “optative of *s*-aorist,” Schumacher & Matzinger 2013: 177). Thus, the optative element *-sh-* is either a direct reflex of the PIE aoristic \**-s-* or an element that developed under the analogical influence of some aorist and imperfect forms. If the former hypothesis is correct, the element *-sh-* (< PIE \**-s-* + \**-ī-*) must have initially evolved in the verbs with the *-v-* aorist (this form itself probably developed from the PIE sigmatic aorist) and later extended to other conjugation classes by analogy (Schumacher & Matzinger 2013: 180–182). The element *-f-* in the verbs ending in a vowel is a result of the devoicing of the Albanian aorist element *-v-* (in the verbs ending in a consonant the *-sh-* marker is attached directly to the stem). As for the endings, they could have been influenced by those of the aorist and imperfect indicative active.

The form 2 sg. SA *-sh-* < PIE \**-s-īs* (Matzinger 2006: 116); 3 sg. SA *-tě* could have evolved under the influence of the 3 sg. present subjunctive active (Matzinger 2006: 116).

### Imperative

2 sg. SA -Ø in verbs belonging to all conjugational classes, -*ë* in some verbs with consonantal stems, -*j* in some verbs with vocalic stems. The stem of verbs with vocalic and consonantal alternations demonstrate *i*-vocalism and the final palatal consonant. -*ë* < PIE \*-e (?).

The 2 pl. imperative form is identical to the corresponding indicative form. An object clitic pronoun can be inserted between the stem and the ending: *dëgjo-më-ni* hear-I. ACC.-2PL.IMP. ‘hear (pl.) me!’; see “Syntax”.

### Participles

The participles in SA are formed by means of several derivational suffixes: -*r(ë)* / Gheg -*n(ë)*; -*ur* / Gheg -*un*; -*në*. Their distribution is determined by the verb class and the stem auslaut; for the majority of verbs, participles are derived from aorist stems. Etymologically all these participle suffixes result from the development of the PIE suffix \*-no-. In some (especially Gheg) dialects and in the old texts there are also participles formed with the formants -*m* (< PIE \*-mo-) and -*të* (< PIE \*-to-). One and the same verb can often use different suffixes of the participle in different dialects or in old texts written by different writers, and sometimes even in the texts of the same writer. Semantic differences between participles formed with different suffixes are not registered. On the short forms of particles in modern Gheg dialects, see above (“Basic Patterns of Dialectal Variation”).

## WORD FORMATION

This section is mainly based on Xhuvani and Çabej 1976a, 1976b, Camaj 1966, Orel 2000: 151–177; Matzinger 2006: 137–140, *in print*; Topalli 2011: 1149–1175; see also Schumacher and Matzinger 2013: 188–190, Sasse 1991: 233–267.

The main word-formation means in Albanian are affixation, compounding and conversion.

### Affixation

Albanian uses both suffixation and prefixation; as in other Indo-European languages suffixation prevails in Albanian.

#### *Suffixation*

A. *Nouns*. There are several chronological layers in the nominal word-formation suffixes of Albanian. First, there are suffixes that can be traced back to the PIE epoch (see Camaj 1966: 113–124): -*k-/-g-* (SA *petk* ‘clothes’, cf. *petë* ‘thin layer of dough’), -*p-/-b-* (SA *thelb* ‘kernel of a nut’), -*l-/-r-* (*kërmill* ‘snail, slug’). As a rule, these suffixes are unproductive in modern Albanian, and their derivational meaning is not transparent.

Then there are many suffixes that are productive or semi-productive in modern Albanian. They derive Albanian nouns from nouns, adjectives and verbs; some of these suffixes have been borrowed from Latin, Slavic and Turkish. Among the most widespread nominal suffixes are the following.

- Suffixes of abstract nouns. SA *-i* (SA *liri* ‘freedom’ from SA (*i, e*) *lirë* ‘free’); this suffix is used to derive nouns from nouns, verbs and adjectives. It goes back to PIE *\*-iyeh<sub>2</sub>*. SA *-im*, cf. SA *pushim* ‘rest, pl. recreation’, from SA *pushoj* ‘to rest’, derives nouns from verbs and goes back to PIE *\*-imu* (see Klingenschmitt 1994: 225); etc.
- Suffixes of nomina agentis. SA *-(ë)s* and its variants *-es, -ues, onjës* (*ndihmës* ‘assistant’ from *ndihmoj* ‘to help’) is used to derive nouns (and adjectives) from nouns and verbs < PIE *\*-ik<sup>w</sup>yo-* (cf. Arm. *-ič<sup>c</sup>-*). SA *-tor* (with variants *-tuar, -tuer, -tar, -or* (*-uer, -uar, -ar* (SA *sherbëtor* ‘attendant’ from SA *sherbëj* ‘to serve’); in all probability all these suffixes are borrowed from Latin; they are used to derive nouns and adjectives from both nouns and verbs. There are also some suffixes of Slavic origin; they can be used to derive not only nomina agentis in a narrow sense but also other nouns that refer to people and sometimes have a pejorative semantic shade: *-ac/-ec, -ash, -avec, -nik*; cf. SA *mburravec* ‘boaster’ from *mburr* ‘to brag’.
- Feminine suffixes: *-eshë, ushë* (SA *luaneshë* ‘lioness’ from SA *luan* ‘lion’), < Lat. *-issa; onjë* < Lat. *-onia* (?); and a borrowed Slavic suffix *-icë*.
- Diminutive suffixes: *-th* (for nouns of masculine gender, < PIE *\*-ko-*), *-zë* (for nouns of feminine gender, < PIE *\*-dyā*), cf. SA *birth* ‘sonny’ from SA *bir* ‘son’. In some Albanian dialects, especially in Arbresh, the use of diminutives is abundant: diminutive suffixes are used there not only with nouns but also with adjectives, adverbs and even verbs.
- Suffixes with the collective and locative meaning: *-shtë* (SA *vreshtë, Gheg vneshtë*(ë) ‘vineyard’ from SA *verë, Gheg venë* ‘vine’) < *\*-V-st-* (this suffix is encountered in some old Balkan geographic names).

B. *Adjectives*. The adjectives share some derivational suffixes with nouns (see above); this is characteristic of unarticulated adjectives. There are some suffixes typical of articulated adjectives:

- *-m(ë)*, cf. SA (*i, e*) *sotëm/me* ‘today’s’ from *sot* ‘today’, is used to derive adjectives from words belonging to various parts of speech, < PIE *\*-mo-*.
- *-shëm*, SA (*i, e*) *kujdesshëm/me* ‘careful’, from *kujdes* ‘care’, is used to derive adjectives from nouns, adverbs and verbs (the latter possibility is attested only in the Gheg area), < *\*-sh* (ablative ending?) + *\*-mo-* (see above).
- *-t(ë)*, cf. SA (*i, e*) *hekurt* ‘iron (adj.)’, from SA *hekur* ‘iron’, is used to derive adjectives from adverbs, verbs and nouns, < PIE *\*-to-*.

C. *Adverbs*. A large part of Albanian adverbs are formed by means of conversion (see below). There are, however, some suffixes that are used to derive adverbs from other parts of speech: *-isht* (SA *besnikërisht* ‘faithfully’ from SA *besnik* ‘faithful’, < PIE *\*-isto-*); *-thi* (SA *fluturimthi* ‘as in the flight’, from SA *fluturim* ‘flight’); *-as/-azi* (SA *gjunjas* ‘on one’s knees’ from SA *gjunjë* ‘knee’) and some others.

D. *Verbs*. On the old IE verbal stem extension in Albanian see above (“Morphology, Verb system: the origin”).

### Prefixation

In modern Albanian, prefixation is characteristic of both the nominal and verbal system. Many prefixes are homonyms to adverbs and prepositions: SA *mbijetoj* ‘to survive’ from *mbi* ‘over’ + *jetoj* ‘to live’; SA *nënçmim* ‘underestimation’ from *nën* ‘under’ + *çmtoj* ‘estimate’. Some prefixes in SA are borrowed: SA *riarmatos* ‘to reweapon’ from *ri-* (from Ital. *ri-*) + *armatos* ‘to weapon’. There are, however, preverbs that form an older chronological

layer. Many such preverbs (see the list in Schumacher & Matzinger 2013: 188–190; also Camaj 1966: 103–113) have lost their meaning in the modern language: cf. e.g. SA *v-desh* ‘to die’ and *des* in texts by old writers. Combinations of two preverbs are encountered very rarely: SA *sh-për-blej* ‘reward’, cf. SA *blej* ‘buy’.

*Compounding* is widely represented in modern Albanian. There are various types of compounds, for example: SA *deledash* ‘hermaphrodite’ from *dele* ‘sheep’ + *dash* ‘ram’ (*dvandva*); SA *ditëlindje* ‘birthday’ from *ditë* ‘day’ + *lindje* ‘birth’ (*tatpuruṣa*); SA *lulëkuqe* ‘poppy’ < *lule* ‘flower’ + (*i, e*) *kuqe* ‘red’ (*karmādhara*); SA *zemërgur* ‘heartless’ from *zemër* ‘heart’ + *gur* ‘stone’ (*bahuvrihi*). The order of the components within compounds is parallel to word order in the syntax. Compounding in the verbal domain is rarer; see, for example, SA *udhëheq* ‘to direct’ from *udhë* ‘way’ + *heq* ‘to remove’.

There are also compounds with opaque structure, e.g. SA *kushtrim* ‘appeal’ from *kush* ‘who?’ and *trim* ‘brave’.

### Conversion

Two clear patterns of conversion are represented by the processes of adjectivization of participles and substantivization of articulated adjectives: SA *folur* ‘to speak (participle)’ → (*i, e*) *folur* ‘colloquial’ → (*të*) *folurit* ‘speech’. Conversion is also characteristic for the “adjective–adverb” pairs: SA (*i, e*) *mirë* ‘good (adjective)’ – *mirë* ‘well (adverb)’, but it is difficult to determine the direction of the derivation in this case.

### SYNTAX

This section focuses on some of those features of the Albanian syntax that are interesting from an areal and typological point of view. Some remarks on Albanian morphosyntactic characteristics can be found above (see “Morphology”). Albanian syntax is very poorly described from a historical point of view (see, however, Joseph *in print*, Matzinger 2006: 140–144, Topalli 2011: 1179–1281).

Albanian is a nominative-accusative language that codes this alignment by case marking, verb agreement and word order (on the additional marking of objects by means of clitics, see below). Albanian is a language with a relatively free word order (as far as major constituents are concerned). Nevertheless, SVO is clearly the basic word order pattern. Numerous deviations from the SVO pattern are accounted for by the topicalization of the object and by other properties of information structure, by actionality characteristics of the verb, as well as by emphatic and stylistic aspects. Albanian is a pronounced pro-drop language.

In general, Albanian syntax does not differ significantly from the syntax of other modern European languages: Albanian shares the majority of diagnostic features determined for the Standard Average European (SAE) languages (Haspelmath 2001: 1505) (these diagnostic features are mostly syntactic). Some important syntactic traits are shared by Albanian with other languages of the Balkan Sprachbund (see below).

*Syntax of the noun phrase.* The main characteristics of the Albanian noun phrase were described above (“Morphology”). Probably the most striking phenomenon is the presence of a special linking element that agrees with the head noun and is placed both before the adnominal genitive (this pattern is thus a kind of *Suffixaufnahme*) and before adjective-like words. This peculiarity is also found in Balkan Romance (first of all, in Daco-Romanian), which is one of the main pieces of evidence pointing toward intensive

(Proto-)Albanian – (Proto-)Rumanian contacts. A feature that Albanian shares with both Romanian and Bulgaro-Macedonian is a rule according to which the first component of the noun phrase is marked for case and other categories if this position is occupied by an adjective, ordinal numeral, possessive pronoun or another element whose default position is after the head noun, cf.: *djal-i i mirë* boy-NOM.M.SG.DEF. ART.NOM.M.SG. good ‘good boy’ (unmarked word order) and *i mir-i djalë* ART.NOM.M.SG. good-NOM.M.SG. DEF. boy ‘good boy’. In old texts there are examples where both members of the noun phrase get inflections; cf. an example from Buzuk: *vepra-vet mira-ve* act.PL.-GEN./DAT. DEF. good.F.PL.-GEN./DAT. ‘good acts (dat.)’.

Along with postposed nominal dependents, described above, there are some nominal dependents that are found in preposed preposition: cardinal numerals, demonstratives, indefinite and interrogative pronouns and the indefinite article. Some of them are indeclinable when used in the noun phrase (cardinal numerals, some indefinite pronouns, indefinite article); thus, in corresponding noun phrases it is only the noun that declines: *ndonjë njeri-u* some person-GEN./DAT.M.SG.INDEF. ‘(of/to) some person’. The interrogative pronoun (*i, e*) *cilila* ‘what’ behaves as an adjective: *i cil-i njëri?* ART.NOM.M.SG. what-NOM.M.SG.DEF. person ‘what person?’ In noun phrases with demonstrative pronouns both members of the construction decline, and the noun can be either in the indefinite or in the definite form: *këtij njeri-u* this.GEN./DAT.M.SG. person-GEN./DAT.M.SG. INDEF. vs. *këtij njeri-ut* this.GEN./DAT.M.SG. person-GEN./DAT.M.SG.DEF. ‘of/to this person’.

*Ablative construction.* Albanian has a special construction in which the attribute is expressed by an ablative form in the singular or in the plural and has a non-referential, qualifying meaning: *këngë dashuri-e* song love-ABL.F.SING.INDEF. ‘love song’, *leng fruta-sh* juice fruit.PL.-ABL.PL.INDEF. ‘fruit juice’. In this construction, both the head and the dependent must be used in the indefinite form; this means that when the head noun is in the definite form, the attribute should be expressed by the genitive, while the ablative construction is ruled out. When the ablative form is in the plural, the ending *-sh* (and not *-ve*) must be used.

*Prepositional constructions.* Albanian prepositions typically govern nouns in the accusative and ablative cases; there are also some prepositions that evolved through grammaticalization of prepositional phrases, which govern the genitive: *në rast të* ‘in (the) case of’. The majority of prepositions govern one case. The exceptions are the prepositions *për* ‘for, about’ and *ndaj* ‘toward, by’, which govern both the accusative and ablative cases. Location at some point and motion towards that point are grammatically expressed by identical prepositional constructions in Albanian, which is also typical of other Balkan languages: *në shtëpi* [in house] ‘at home, to the house’. Albanian has also two prepositions, *nga* ‘from’ and *te(k)* ‘to’, which govern the nominative.

Albanian belongs to the languages with transitive *have-possessive constructions*. *Existential clauses* are formed with the 3 sg. form of the verb *kam* ‘to have’: *Këtu ka disa studente* Here have.3SG.PRES. some students. ‘There are some students here’. A construction with *jam* ‘to be’ can also have an existential meaning: *Janë ca njerëz, që . . .* be.3PL.PRES. some people, which . . . ‘There are some people, who. . .’.

*Weak (clitic) pronouns.* These pronouns double the direct and indirect objects. They may be the only means of expression of direct or indirect objects in clauses with 1st or 2nd person objects or in clauses where the weak 3rd person pronoun is used in an anaphoric function. On the form of weak pronouns, see above (“Morphology”). Dative weak pronouns are obligatory in SA; they double not only “true” indirect objects but

all clause-level dependents in the dative case, including nominal dative forms that are used in adverbial functions. In the old texts and in the dialects dative weak pronouns are not consistently obligatory. Pronominal doubling of direct objects is optional, except for direct objects expressed by the pronouns *i gjithë* ‘all’ and *i tërë* ‘all’, which are obligatorily doubled. Direct object doubling depends on information-structural properties of the clause: rhematic direct objects are never doubled, cf. (1) and (2):

- 1) *Agim-in* *e* *gjeta* *në* *dhomë*  
 Agim-ACC.M.SG.DEF. PRO.ACC.3SG I.found in room  
 ‘I found Agim in the room’.
- 2) *Në dhomë gjeta* *Anil-ën*  
 in room I.found Anila-ACC.F.SG.DEF.

‘In the room I found Anila (It was Anila whom I found in the room)’.

Accusative forms that are not direct objects are never doubled.

Weak pronouns are placed immediately before the verb. The only exception are imperative constructions: weak pronouns can be placed after the 2 sg. form and can be incorporated within the 2 pl. form (see “Morphology” above). When both a dative and an accusative weak pronoun are used in the same clause, the order is always “dat. + acc.” As a result, there is a series of coalesced forms: *ma* (*më + e*) ‘1 sg. dat. + 3 sg. acc.’, *ta* (*të + e*) ‘2 sg. dat. + 3 sg. acc.’, *ia* (*i + e*) ‘3 sg. dat. + 3 sg. acc.’, etc:

- 3) *Ia* *dhashë* *libr-in*  
 PRO.3SG.DAT.PRO.3SG.ACC. I.gave book-ACC.M.SG.DEF.  
*mesues-it*  
 teacher- DAT.M.SG.DEF.

‘I gave the book to the teacher’.

Pronominal doubling of objects is a characteristic trait of the Balkan Sprachbund. The form of the coalesced pronouns suggests the PA period as the period when they evolved and thus again the period of (proto)Albanian-(proto)Romanian contacts.

*Negation.* Albanian has different negation particles for the real (indicative, admirative) and unreal (subjunctive, optative, imperative) moods; it thus preserves the PIE distinction, which is similarly preserved in Indo-Iranian, Armenian and Greek. In real contexts SA *s* (for clause negation also a longer variant *as* < PIE \*(ne) h<sub>2</sub>oyu-k<sup>w</sup>id) or *nuk* (< PIE \*ne-k<sup>w</sup>e) is used; in unreal contexts SA *mos* is used (< PIE \*mē-k<sup>w</sup>e).

*Complementation.* There are two main models of complementation in SA. One pattern employs conjunctions *se* ‘that’ (< PIE \*k<sup>w</sup>e) or *që* ‘that, what’ (< PIE \*k<sup>w</sup>i/o-); *se* is preferred in contexts with most semantic types of matrix verbs. Another pattern involves the use of the subjunctive in the dependent clause (see “Morphology” above). The choice between these two models is determined by (1) whether the event in the dependent clause is factual (real) or non-factual (non-real) and (2) the degree of boundedness of the events in the main and subordinate clauses (see Matras 2002: 179–185). Thus, there is a kind of continuum stretching from epistemic verbs and perception verbs on one pole of the continuum (these verbs favour the conjunction pattern; cf. (4)) to modal and phasal verbs on the other pole of the continuum (these verbs favour the subjunctive pattern; cf. (5)).

The conjunction *që* can introduce a dependent clause with a subjunctive verb form (for example, if the subjects of the main and dependent clauses are not coreferential) (6).

- 4) *Ai e di se keni*  
 He PRO.3SG.ACC. he.knows that have.2PL.PRES.  
*dobësi*  
 weakness  
 ‘He knows that you have weakness’.

- 5) *Dua të shkoj*  
 Want.1SG.PRES. SUBJ go.1SG.PRES.  
 ‘I want to go’.

- 6) *Dua që të shkosh*  
 Want.1SG.PRES. that SUBJ. go.2SG.PRES.  
 ‘I want that you go’.

Such a two-fold complementation strategy is a characteristic trait of the Balkan languages. In Gheg dialects (including old Gheg texts) the subjunctive strategy competes with a strategy with the infinitive of the “*me* + participle” type (see Arapi 2010).

*Conditional sentences.* There is a great deal of variation in the ways Albanian can express conditional relations, especially in the protasis, where the choice of the verb form is determined by (1) the choice among several conjunctions (*po, në, nëse, sikur*) or the absence of the conjunction; and (2) the type of conditional meaning: real or potential vs. irreal. The indicative, optative, admirative and subjunctive forms are used in the case of a real or potential condition, whereas the imperfect or the pluperfect of subjunctive (or the subjunctive of admirative) is used in the case of an irreal condition. In the apodosis of the irreal conditional sentences the imperfect and pluperfect indicative as well as the present and past conditional can be used.

*Relative sentences.* SA has the inflected relativizer (*i, e cili* ‘which, who’ (7), which can be used for relativization of any syntactic position in the dependent clause. This pronoun is ousting the old general relativiser *që/çë*, which can be used in modern SA exclusively for the relativization of subjects and direct objects (8).

- 7) *ky trim, për të cilin*  
 this brave.man on ART.ACC.SG. which.ACC.M.SG.DEF.  
*tregoheshin legjenda, trokiti mu në derën*  
 tell.3PL.IMP.F.PASS. tales he.knocked just in door.ACC.F.DEF.  
*e tyre*  
 ART.ACC.SG.DEF. their  
 ‘This brave man, about whom the tales were told, knocked directly at their door’.

- 8) *Njerëzit që bënin sehir nuk nxitonin*  
 people which do.3PL.IMP. looking not hurry.3PL.IMP.F.  
 ‘People, which looked around, did not hurry’.

## FURTHER READING

*Bibliographic data* on Albanian and Albanian linguistics are published regularly in “Bibliographie linguistique” and “Indogermanische Chronik” (before 1949 Albanian linguistics was reviewed in *Indogermanisches Jahrbuch*); see also Schmitt-Brandt 1968, Hamp 1972, Daka, 1975, 1984, Ködderitzsch 1994.

*Overviews of the history of the albanological research* can be found in Jokl 1917, Schmitt-Brandt 1968, Hamp 1972, Ködderitzsch 1994. There are two substantial introductions to Albanian studies: Fiedler 2006 and Matzinger 2006; see also an overview of the main problems of Albanian language history in Sh. Demiraj 2013 (with a detailed English summary).

The basic *monolingual dictionaries* are Kostallari 1980 and Thomai 2006 (see also Cipo et al. 1954). *Albanian-English dictionaries* include Mann 1948 (a “historical” dictionary with very rich dialect material) and Newmark 1999. The annotated Albanian corpus (19 million tokens) is now available at <http://web-corpora.net/AlbanianCorpus/search/>.

The main *etymological dictionaries* are Meyer 1891 (the first Albanian etymological dictionary), Huld 1984 (basic IE etymologies of Albanian), B. Demiraj 1997 (it encompasses only the IE part of the Albanian lexicon), Orel 1998 (IE etymons and early borrowings), and Cabej 1976–2014 (this *opus magnum* by the greatest Albanian linguist encompasses, unfortunately, only a selection of the Albanian lexicon there is a French summary); see also the etymological dictionary of Albanian verbs of PIE descent in Schumacher and Matzinger 2013: 965.

The best Albanian *reference grammar* is Buchholtz and Fiedler 1987; see also Newmark, Hubbard and Prifti 1982 (based in many parts on Domi 1976) and Camaj 1984.

Among the *historical grammars* Mann 1977 and Orel 2000 employ a comparative approach, whereas Sh. Demiraj 1986 is the most detailed collection of Old Albanian and dialectal language material (see also Sh. Demiraj 1993); Schumacher and Matzinger 2013 contains a thorough analysis of the history of the Albanian verb system.

## NOTE

- 1 This reconstructed PA stage may actually reflect a somewhat later stage in the development of the Albanian language. Indeed, on the one hand, there are some grounds to believe that (Proto)Gheg and (Proto)Tosk dialects had intensive contact in the first half of the second millennium AD. However, on the other hand, some similar processes of language change may have developed in both Gheg and Tosk independently. These circumstances can influence the assumed reconstruction.

## REFERENCES

- Arapi, Inna 2010, *Der Gebrauch von Infinitiv und Konjunktiv im Altalbanischen mit Ausblick auf das Rumänische*, Kovač, Hamburg
- Asenova, Petza 2002, *Balkansko jezikoznanije. Osnovni problemi na balkanskija jezиков ѕъјуз*, Faber, Sofja
- Bonnet, Guillaume 1998, *Les mots latins de l'albanais*, L'Harmattan, Paris
- Bopp, Franz 1854, “Über das Albanesische in seinen verwandschaftlichen Beziehungen,” *Kön. Preuss. Akad. D. Wiss. Berlin Abh. Philol. Hist. Kl.*: 459–549



- Boretzky, Norbert 1975–1976, *Der türkische Einfluss auf das Albanische*, I–II (=Albanische Forschungen 11–12), Harrassowitz, Wiesbaden
- Boretzky, Norbert 1977, “Die Morphologie des albanischen Verbums,” *Zeitschrift für Balkanologie* B. 13: 20–56
- Breu, Walter 2010, “Mood in Albanian,” in: Rothstein, Björn & Thieroff, Rolf (eds.), *Mood in the Languages of Europe*, John Benjamins, Amsterdam–Philadelphia: 447–472
- Breu, Walter 2013, “The category of mood in albanian and its varieties,” *Hylli i Dritës* 33/1–2: 15–50
- Buchholz, Oda & Fiedler, Wilfried 1987, *Albanische Grammatik*, VEB Verlag Enzyklopädie, Leipzig
- Çabej, Eqrem 1976, “Fonetika historike,” in: Eqrem Çabej (ed.), *Studime gjuhësore* III, Rilindje, Prishtinë: 105–146
- Çabej, Eqrem 1976–2014, *Studime etimologjike në fushë të shqipes* I–VII, Akademija e shkencave e Republikës së Shqipërisë. Instituti i gjuhësisë dhe i letërsisë, Tiranë
- Camaj, Martin 1966, *Albanische Wortbildung. Die Bildungsweise der älteren Nomina* (=Albanische Forschungen 6), Harrassowitz, Wiesbaden
- Camaj, Martin 1984, *Albanian Grammar with Exercises, Chrestomathy and Glossaries*, Collaborated on and translated by Leonard Fox, Harrassowitz, Wiesbaden
- Cipo, Kostaq, Çabej, Eqrem, Domi, Mahir, Krajni, Anton & Myderrizi, Osman 1954, *Fjalori i gjuhës shqipe*, Instituti i Shkencave, Sekcioni i Gjuhës e i Letërsisë, Tiranë
- Daka, Palok 1975, *Bibliografi e studimeve dhe e artikujve për gjuhën shqipe (1945–1974)*, Akademija e shkencave e RP të Shqipërisë. Instituti i gjuhësisë dhe i letërsisë, Tiranë
- Daka, Palok 1984, *Bibliografi retrospektive e gjuhësisë dhe e onomastikës shqiptare : 1879–1944 : bibliografi me anotacione hartuar sipas të dhënave të shtypit periodik shqiptare mbi Shqipërinë të viteve 1848–1944*, Akademia e shkencave, Tiranë
- de Vaan, Michiel 2004, “PIE \*e in Albanian,” *Die Sprache* 44: 70–83
- de Vaan, Michiel 2014, “Old Albanian presents in -ën,” paper presented on the symposium “*Sprache und Kultur der Albaner*,” 5–7 June 2014, Pogradec (Albania)
- de Vaan, Michiel *in print*, “The phonology of Albanian”
- Demiraj, Bardhyl 1997, *Albanische Etymologien Untersuchungen zum albanischen Erbwortschatz*, Editions Rodopi, Amsterdam–Atlanta, GA
- Demiraj, Bardhyl 2010, “Shqiptar – the generalization of this ethnic name in the XVIII century,” in: Demiraj, Bardhyl (ed.), *Wir sind die Deinen. Studien zur albanischen Sprache, Literatur und Kulturgeschichte* (=Albanische Forschungen 29), Harrassowitz, Wiesbaden: 533–565
- Demiraj, Shaban 1986 *Gramatikë historike e gjuhës shqipe*, Shtëpia botuese “8 nëntori”, Tiranë
- Demiraj, Shaban 1993, *Historische Grammatik der albanischen Sprache*, Österreichische Akademie der Wissenschaften, Wien
- Demiraj, Shaban 1996, *Fonologjia historike e gjuhës shqipe*, Toena, Tiranë
- Demiraj, Shaban 2013, *Gjuha shqipe dhe historia e saj*, Onufri, Tiranë
- Desnickaja, Agnija 1968, *Slavjanskije zainstvovanija v albanskom jazyke*, Izdatel'stvo Akademii nauk SSSR, Moskva
- Desnickaja, Agnija 1976, “Kategorija sobiratelnosti i kategorija massy v istorii grammaticheskogo stroja albanskogo jazyka,” in: Desnickaja, Agnija (ed.), *Grammaticheskij stroj balkanskih jazykov*, Nauka, Leningrad: 5–104
- Desnickaja, Agnija 1984, “Nekotoryje voprosy izuchenija rannih periodov istorii albanskogo jazyka,” in: Desnickaja, Agnija (ed.), *Sravnitelnoje jazykoznanije i istorija jazykov*, Nauka, Leningrad: 5–104

- Domi, Mahir (ed.) 1976, *Fonetika dhe gramatika e gjuhës së sotme letrare shqipe, V, II, Morfologjia*, Akademia e shkencave e RPS të Shqipërisë. Instituti i gjuhësisë dhe i letërsisë, Tiranë
- Fiedler, Wilfried 1981, "Zur Herkunft der albanischen synthetischen Nichtaktivformen," *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 34: 534–543
- Fiedler, Wilfried 2004, *Das albanische Verbalsystem in der Sprache des Gjon Buzuku*, Akademia e shkencave dhe e Arteve e Kosovës, Prishtinë
- Fiedler, Wilfried 2006, *Einführung in die Albanologie*, Meissen (<http://www.linguistik.uni-erlangen.de/~bmkabash/Fiedler/EINF-ALB.20070524.BKTT.pdf>)
- Gjinari, Jorgji; Beci, Bahri; Shkurtaj, Gjovalin & Gosturani, Xheladin 2007, *Atlasi dialektologjik i gjuhës shqipe, V. 1*, Università degli Studi di Napoli L'Orientale, Napoli
- Haarmann, Harald 1972, *Der lateinische Lehnwortschatz im Albanischen*, Buske, Hamburg
- Hamp, Eric P. 1966, "The position of Albanian," in: Birnbaum, Henrick & Puhvel, Jaan (eds.), *Ancient Indo-European Dialects*, University of California Press, Berkeley–Los Angeles: 97–121
- Hamp, Eric P. 1971: 'Fils' et 'fille' en italique: nouvelle contribution, *Bulletin de la Société de Linguistique de Paris* 66: 213–227
- Hamp, Eric P. 1972, "Albanian", in: Sebeok, Thomas (ed.), *Current Trends in Linguistics V.9/2, Linguistics in Western Europe*, Mouton, The Hague–Paris: 1626–1692
- Haspelmath, Martin 2001, "The European linguistic area: Standard Average European," in: Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, Wolfgang Raible (eds.), *Language Typology and Language Universals*, Bd. 2, Mouton de Gruyter, Berlin–New York: 1492–1510
- Hock, Wolfgang 2005, "Zur Vorgeschichte des albanischen Lautsystem," in: Meiser, Gerhard & Matzinger, Joachim (eds.), *Sprachkontakt und Sprachwandel, Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.–23. September 2000, Halle an der Saale*, Reichert Verlag, Wiesbaden: 261–274
- Huld, Martin E. 1984, *Basic Albanian etymologies*, Slavica Publishers, Columbus, Ohio
- Jasanoff, Jay 2012, "Long-vowel preterites in Indo-European," in: Melchert, Craig H. (ed.), *The Indo-European verb, Proceedings of the Conference of the Society for Indo-European Studies, Los Angeles, 13–15 September 2010*, Reichert Verlag, Wiesbaden: 127–135
- Jokl, Norbert 1917, "Albanisch," in: Streitberg, Wilhelm (ed.), *Geschichte der indogermanischen Sprachwissenschaft*, Bd. III: 109–154
- Jokl, Norbert 1924, "Albaner," in: Ebert, Max (ed.), *Reallexikon der Vorgeschichte*, Bd. 1, Walter Gruyter & Co, Berlin: 84–94
- Jokl, Norbert 1934, "Slaven und Albaner," *Slavia* XIII: 281–325
- Jokl, Norbert 1963, "Die verwandtschaftsverhältnisse des Albanischen zu den übrigen indogermanischen Sprachen," *Die Sprache* 9: 113–156
- Joseph, Brian *in print*, "The syntax of Albanian"
- Jusufović, Lumnije 2011, *Die zentralgegische Mundartengruppe in Mazedonien* (Albanische Forschungen 30), Harrasowitz Verlag, Wiesbaden
- Kaluzhskaya, Irina 2001, *Paleobalkanskije relikty v sovremennyh balkanskijh jazykah*, Indrik, Moskva
- Klingenschmitt, Gert 1981, "Albanisch und Indogermanisch," *Münchener Studien zur Sprachwissenschaft* 40: 93–131
- Klingenschmitt, Gert 1994, "Das Albanische als Glied der indogermanischen Sprachfamilie (Tischvorlage)," in: Rasmussen, Jens Elmegård & Nielsen, Benedicte (eds.), *In*

- honorem Holger Pedersen, Kolloquium der Indogermanischen Gesellschaft vom 25. bis 28. März 1993 in Kopenhagen*, Reichert Verlag, Wiesbaden: 221–233
- Ködderitzsch, Rolf 1994, “Albanische Forschungen (speziell ab 1968),” *Kratylos* 39: 1–21
- Kostallari, Androklli (kryered.) 1980, *Fjalor i gjuhës së sotme shqipe*, Akademia e Shkencave të RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë
- Landi, Addolarata 1989, *Gli elementi latini nella lingua albanese*, Edizioni Scientifiche Italiane, Napoli
- La Piana, Marco 1949, *Studi linguistici albanesi*, Pezzino, Palermo
- Mann, Stewart 1948, *An Historical Albanian-English Dictionary*, Longmans, Green and Co, London–New York–Toronto
- Mann, Stewart 1977, *An Albanian Historical Grammar*, Buske, Hamburg
- Matras, Yaron 2002, *Romani. A Linguistic Introduction*, Cambridge University Press, Cambridge
- Matzinger, Joachim 2006, *Der altalbanische Text [E] Mbsuame e Krështerë (Dottrina Cristiana) des Lekë Matrënga von 1592, eine Einführung in die albanische Sprachwissenschaft*, Verlag J. H. Röll, Dettelbach
- Matzinger, Joachim 2009, “Die Albaner als Nachkommen der Illyrier aus der Sicht der historischen Sprachwissenschaft,” in: Schmitt, Oliver Jens & Frantz, Eva (eds.), *Albanische Geschichte. Stand und Perspektiven der Forschung*, R. Oldenburg Verlag, München: 13–35
- Matzinger, Joachim 2010, “Die albanische Sprache im Zeitalter Skanderbegs, Teil 1: Ein Überblick zur vorliterarischen Dokumentation,” in: Genesin, Monica, Matzinger, Joachim & Vallone, Giancarlo (eds.), *The Living Skanderbeg. The Albanian Hero between Myth and History*, Verlag Dr. Kovač: 41–69
- Matzinger, Joachim 2012a, “Zwischensprachen – areallinguistische Bemerkungen aus dem Bereich des Balkanindogermanischen,” in: Sadovski, Velizar & Stifter, David (eds.), *Iranistische und indogermanistische Beiträge in memoriam Jochem Schindler (1944–1994)*, Verlag der Österreichischen Akademie der Wissenschaften, Wien: 137–159
- Matzinger, Joachim 2012b, “Der lateinisch-albanische Sprachkontakt und seine Implikationen für Vorgeschichte des Albanischen und der Albaner,” in: Dahmen, Wolfgang u. a. (eds.), *Südosteuropäische Romania. Siedlungs-/Migrationsgeschichte und Sprachtypologie*, Narr Verlag, Tübingen: 75–103
- Matzinger, Joachim *in print*, “The Lexicon of Albanian”
- Meyer, Gustav 1891, *Etymologisches Wörterbuch des albanesischen Sprache*, Verlag von Karl J. Trübner, Strassburg
- Meyer, Gustav 1892, “Albanesische Studien III: Lautlehre der indogermanischen Bestandteile des Albanesischen,” in: *Sitzungsberichte der phil.-hist. Klasse der kais. Acad. d. Wissenschaften*, Bd.125, 11. Abh., Wien: 1–95
- Mihăescu, Haralambie 1966, “Les éléments latins de la langue albanaise,” *Revue des études sud-est européennes* 4: 5–33, 323–353
- Newmark, Leonard 1999, *Albanian English Dictionary*, Oxford University Press, Oxford
- Newmark, Leonard, Hubbard, Philip & Prifti, Peter 1982, *Standard Albanian. A Reference Grammar for Students*, Stanford University Press, Stanford
- Ölberg, Hermann M. 1972, “Griechisch-albanische Sprachbeziehungen,” in: Muth, Robert (ed.), *Serta philologica aenipontana II*, Innsbrucker Beiträge zur Kulturwissenschaft, Innsbruck: 33–64

- Ölberg, Hermann 2013, *Untersuchungen zum indogermanischen Wortschatz des Albanischen und zur diachronen Phonologie aufgrund des Vokalsystems*, Herausgegeben von Bardhyl Demiraj (=Albanische Forschungen 35), Harrassowitz Verlag, Wiesbaden
- Orel, Vladimir 1998, *Albanian Etymological Dictionary*, Brill, Leiden
- Orel, Vladimir 2000, *A Concise Historical Grammar of the Albanian Language: Reconstruction of Proto-Albanian*, Brill, Leiden
- Pedersen, Holger 1894, "Bidrag til den albanesiske Sproghistorie," in: *Festskrift til Vilhelm Tomsen*, Gyldendal, København: 246–257
- Pedersen, Holger 1895, *Albanesische Texten mit Glossar*, S.Hirzel, Leipzig
- Pedersen, Holger 1897, "Das albanesische Neutrum," *Zeitschrift für vergleichende Sprachforschung* 34: 283–291
- Pedersen, Holger 1900, "Die Gutturale im Albanesischen," *Zeitschrift für vergleichende Sprachforschung* 36: 277–342
- Pisani, Vittore 1950, "L'albanais et les autres langues indoeuropéennes," in: *Mélanges Henri Grégoire II* (Annuaire de l'Institut de philologie et d'histoire orientales et slaves, X), Bruxelles: 519–538
- Pisani, Vittore 1959, "Lexikalische Beziehungen des Albanesischen zu den anderen indogermanischen Sprachen," in: Pisani, Vittore (ed.), *Saggi di linguistica storica*, Rosenberg & Sellier, Torino: 115–136
- Porzig, Walter 1974, *Die Gliederung des indogermanischen Sprachgebiets*<sup>2</sup>, Carl Winter, Heidelberg
- Roques, Mario Louis 1932, *Recherches sur les anciens textes albanais. Avec huit facsimilés*, Paul Geuthner Paris
- Sasse, Hans-Jürgen 1991, *Arvanitika. Die albanischen Sprachreste in Griechenland, Teil 1*, Harrassowitz, Wiesbaden
- Schaller, Wilhelm 1975, *Die Balkansprachen. Eine Einführung in die Balkanphilologie*, Carl Winter, Heidelberg
- Schmitt-Brandt, Robert 1968, "Albanologische Forschungen (speziell ab 1958)," *Kratylos* 13: 1–26
- Schumacher, Stephen & Matzinger, Joachim 2013, *Die Verben des Altalbanischen Belegwörterbuch, Vorgeschichte und Etymologie*, Unter Mitarbeit von Anna-Maria Adaktylos (=Albanische Forschungen 33), Harrassowitz, Wiesbaden
- Selishchev, Afanasij M. 1931, *Slavjanskoje naselenije v Albanii*, Izdanije Makedonskogo nauchnogo instituta, Sofia
- Shuteriqi, Dhimitër S. 1976, *Shkrimët shqipe në vitet 1332–1850*, Akademia e Shkencave të RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Tiranë
- Svane, Gunnar 1992, *Slavische Lehnwörter im Albanischen*, Aarhus University Press, Aarhus
- Thomai, Jani (red. shk.) 2006, *Fjalor i gjuhës shqipe*, Akademia e Shkencave e Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Tiranë
- Topalli, Kolec 2006, *Phonetika historike e gjuhës shqipe*, Dituria, Tiranë
- Topalli, Kolec 2011, *Gramatikë historike e gjuhës shqipe*, Kristalina–KH, Tiranë
- Vătăşescu, Cătălina 1997, *Vocabularul de origine latină în limba albaneză în comparație cu româna*, Institutul de Tracologie, Bucureşti
- Vermeer, Willem 2008, "The prehistory of the Albanian vowel system," in: Lubotsky, Alexander; Shaeken, Jos & Wiedenhof, Jeroen (eds.), *Evidence and counter-evidence, Essays in honour of Frederik Kortlandt, vol.1* (= *Studies in Slavic and General Linguistics* 32), Rodopi, Amsterdam–New York: 591–608

- Weigand, Gustav 1927, "Sind die Albaner die Nachkommen der Illyrer oder der Thraker?," *Shuteriqi* in: *Balkan-Archiv* III: 227–251
- Xhuvani, Aleksander & Çabej, Eqrem 1976a, "Parashtesat e gjuhës shqipe," in: Çabej, Eqrem (ed.), *Studime gjuhësore*, V. III, Rilindje, Prishtinë: 49–188
- Xhuvani, Aleksander & Çabej, Eqrem 1976b, "Prapashtesat e gjuhës shqipe," in: Çabej, Eqrem (ed.), *Studime gjuhësore*, V. III, Rilindje, Prishtinë: 191–307
- Ylli, Xhelal 1997, *Das slavische Lehngut im Albanischen. 1. Teil. Lehnwörter*, Otto Sagner, München

---

# LANGUAGE INDEX

---

- Abkhaz 155–6, 161  
Aeolic Greek (Aeol.) 4, 287–8, 290, 292, 307;  
*see also* Greek  
Afro-Asiatic 7, 153, 155–6, 160–2, 164, 166–7  
Ainu 153  
Akkadian 3, 165  
Albanian (Alb.) 3, 5, 13, 422, 427, 455–6, 489,  
529, 544, 558; adjectives 580–2, 598; adverbs  
555, 580–2, 598; Arabic borrowings 558;  
areal position of 558; Balkanization of 558;  
Cham dialect 552, 567; current geographic  
distribution 552; dialects 552–3, 567;  
documentary evidence 553–4; ethnic and  
language name 554–5; Europeanization of 558;  
further reading 603; Greek borrowings 557–8;  
internationalisms 558; Latin borrowings  
557, 572, 597; layers of the lexicon 557–9;  
map 553; Modern 553, 558, 560; numerals  
584–5; Old 558, 560; origin and early history  
555–7; Persian borrowings 558; pre-Proto-  
Albanian period 559–60; pronouns 582–4;  
Proto- 557–60; Slavic borrowings 557, 597;  
stress 572–3; Turkish borrowings 558, 597;  
word formation 597–9; *see also* Albanian  
morphology; Albanian phonology; Albanian  
syntax; Albanian verbs; Albanian word  
formation; Gheg dialect; Tosk dialect  
Albanian morphology: dialectal differences  
(nominal) 575–6; gender assignment of nouns  
575; nominal system (definite declension)  
579–80; nominal system (indefinite declension)  
576–9; noun declensions 574–5; nouns 573–4  
Albanian phonology 560; diphthongization  
565; morphophonological alternations 573;  
obstruents 568–72; sonorants 566–8; stops  
15, 17–19; stress 572; umlaut 564; velars 21;  
vowels 561–6  
Albanian syntax 545; ablative construction 600;  
complementation 601–2; conditional sentences  
602; negation 601; noun phrase 599–600;  
prepositional constructions 600; relative  
sentences 602; weak (clitic) pronouns 600–1  
Albanian verbs: active indicatives 594; aorist  
stems 595–6; basic patterns of dialectal  
variation 592–3; conjugational classes 588–9;  
conjugations of irregular verbs 591–2; general  
information 585–7; imperative 597; imperfect  
active 594–5; imperfect non-active 595; major  
types of aorist stems 590; major vocalic and  
consonantal alterations 590–1; non-active  
indicatives 594; non-finite forms 588; optative  
mood 596–7; participles 597; present stems  
593–4  
Albanian word formation: affixation 597–8;  
prefixation 598–9; suffixation 597–8  
Altaic 7, 153, 155–6, 158–60, 162, 164, 166–7,  
208–9, 211; *see also* Turkic  
American English 2, 400; *see also* English  
Anatolian (Anat.) 3, 6, 159, 287, 428; accent 177;  
adverbs and adpositions 190–1; consonants  
176–7; further reading 194; lexicon 193–4;  
morphosyntax 191; nominal derivation 180;  
nominal inflection 178–80; phonology 15–18,  
21, 176–7; pronouns 180–4; subordinate  
clauses 193; word order 191–3; *see also*  
Anatolian verbs; Proto-Anatolian  
Anatolian subgroup: Carian 3, 172, 175; Hittite  
3, 13, 15–17, 19, 21, 171–3, 178, 357; Luwian  
13, 21, 171–4; Lycian 3, 21, 172, 174–5;  
Lydian 2, 172, 175; maps 173–4; overview  
171–2; Palaic 3, 172–3; Pisidian 3, 175;  
Sidetic 3, 175  
Anatolian verbs: about 184–5; active inflection  
185–8; medio-passive inflection 188; non-finite  
forms 190; verb stem formation 189  
Ancient Greek 13, 21, 33, 49, 55, 120, 134, 145,  
147, 155, 165, 238, 317, 321, 323; *see also*  
Greek; Homeric  
ancient Indo-European languages (AIELs) 111,  
171, 331, 447, 555  
Apabhramśa Prakrits 215, 251  
Aquitanian 356  
Arabic 1, 18; Albanian borrowings from 558  
Arbresh dialect 552, 558, 567, 592–3  
Arcadian (Arc.) 294, 302, 304  
Arcado-Cypriot Greek (AC) 4, 287, 301, 308;  
*see also* Greek  
Armenian (Arm.) 3, 5, 6, 389, 545, 557–8; about  
421–3; Classical 13, 421; consonantal shifts  
18, 20; Eastern 421; further reading 447–8;  
inherited suffixes and clusters 446; Middle 421;  
nominal compounds 446–7; nominal word  
formation 445–7; syntax 447; transcription of  
423; *see also* Armenian morphology; Armenian  
nouns; Armenian phonology  
Armenian morphology: adjectives 438; gradation  
438; *l*-stems 437; nouns 434–8; *n*-stems  
436–7; numerals 440–1; pronouns 438–40;  
*r*-stems 437; verbs 441–5  
Armenian nouns: alternating stem classes 436;  
gradation 438; heteroclitics and irregular  
paradigms 437–8; *l*-stems 437; *n*-stem *harsn*

- 'bride' 436; *n*-stems 436; paradigm of *hayr* 'father' 437; *r*-stems 437; vocalic stem classes 434–6
- Armenian phonology 15–19, 21, 423–34;  
 consonant clusters 431–3; consonants 426–34;  
 diphthongs 424–5; initial position before  
 vowels (\*HV) 430; interconsonantal position  
 (\*CHC) 430; laryngeals 429–30; metathesis  
 432; position after *i/u* before a consonant  
 (\*i/uHC) 430–1; position between vocalic  
 sonant and stop/s (\*RHC) 430; semivowels  
 429; vowels 423–6; *w*-clusters 433–4;  
*y*-clusters 433
- Aromanian 545
- Arvanitika (Arvantica) dialect 552, 554, 558,  
 567, 592
- Aryan 3; *see also* Common Indo-Iranian;  
 Indo-Aryan (IA)
- Assamese 210, 217, 250, 257
- Atharva-Veda (AV) 246, 249
- Attic Greek (Att.) 291, 294, 296–300, 304, 307,  
 314, 317
- Attic-Ionic Greek (AI) 4, 287, 289, 302, 306–7,  
 313; *see also* Proto-Attic-Ionic Greek
- Aukštatic dialect 488–9, 496
- Australian English 2; *see also* English
- Austro-Asiatic 155
- Avestan (Av.) 4, 13, 18, 21; Late 13, 15–17, 19,  
 283; *see also* Gatha-Avestan
- Azeri 1
- Bactrian 266
- Balkan Indo-European 558; *see also*  
 Indo-European
- Balkan Romance 558, 579, 599; *see also*  
 Romance
- Balkan Slavic 544, 579; *see also* Slavic
- Balkan Turkish 1; *see also* Turkish
- Baltic (Balt.) 5, 389, 454, 499, 503, 521, 522,  
 529, 558; about 486–7; adpositions 514;  
 early texts 487–8; East Baltic dialects 488–9;  
 extinct languages 486; further reading 515;  
 map of Baltic tribes 487; *ruki* rule 497; syntax  
 512–15; verbs 486–7; Wackernagel's Law  
 514–15; word order 513–14; *see also* Baltic  
 morphology; Baltic nouns; Baltic phonology;  
 Baltic prosody; Baltic verbs; Balto-Slavic;  
 Common Baltic; East Baltic; Latvian;  
 Lithuanian; Old Prussian; Proto-Baltic; West  
 Baltic
- Baltic Finnic 389, 486, 494, 498, 516n6; *see also*  
 Finnish/Finnic
- Baltic morphology: adjectives 503–5; nouns  
 499–503; numerals 505–6; pronouns 506–8;  
 verbs 508–12
- Baltic nouns 499–503; declension of *a*-stem  
 nouns 502; nominal derivation 503; nominal  
 inflection 499; nominal stem types 500–1;  
 sample *o*-stem case endings 501–2
- Baltic phonology: accent paradigm  
 correspondences 526; accentual paradigms 527;  
 consonant + \**r* sequences 498–9; consonants  
 496–9; palatovelars 498; vowels 489–91
- Baltic prosody: Endzelin's Law 494–5;  
 Leskien's Law 493; Saussure's Law 493; tonal  
 correspondences 495–6; tonal oppositions in  
 Latvian 494; tonal oppositions in Lithuanian  
 492–3; tonal oppositions in Old Prussian 495
- Baltic verbs: aspect 512; athematic presents  
 509; diathesis and ablaut 510; mood 510–11;  
 participles 511–12; preterit 510; semi-thematic  
 verb type 509–10; thematic verb type 509;  
 verbal derivation 512; verbal inflection 508
- Balto-Slavic (BSL) 5, 6, 389, 428, 455, 6,  
 479–83, 496, 506; accentology 481–2, 486n5,  
 486n7, 486n8; circumflex tone 479; palatalized  
 velars 21; voiced aspirated stops 18, 19;  
*see also* Baltic; Proto-Balto-Slavic; Slavic
- Basque 1, 157, 160, 164–5, 356
- Batsbi (Tsova-Tush) 164
- Bavarian 397
- Belorussian/Belarusian (BRuss.) 5, 519, 524, 547;  
 syntax 545; *see also* Slavic
- Bengali 1, 4, 214, 215, 217
- Boeotian Greek (Boe.) 287, 293–5; *see also*  
 Greek
- Bosnian 519; *see also* Slavic
- Bosnian/Croatian/Montenegrin/Serbian  
 (BCMS = Serbo-Croatian) 5, 527; dialects 557;  
 phonology 16–17
- Brabantian 397
- Breton (Bret.) 4, 352, 361, 364, 378; *see also* Old  
 Breton; Primitive Breton
- British 2, 352, 358, 364–5, 367, 389; *see also*  
 Brythonic; English
- Brythonic (Brittonic/British) 4, 352–3, 357–8,  
 362–3, 367–8, 370, 375–6; Gallo- 358;  
 phonology 364–5; vocalism 364–5
- Bulgarian (Bulg.) 5, 519, 522, 543–4, 546;  
 dialects 557; syntax 545; *see also* Slavic
- Bulgaro-Macedonian 558, 600; *see also*  
 Bulgarian; Macedonian
- Burgund 5
- Burmese, Inscriptional 389
- Burushaski 156, 164, 208
- Byzantine/Medieval Greek 4
- Čakavian Croatian (Čak.) 527, 538, 544
- Cappadocian 287
- Carian 3, 172, 175
- Catalan 4
- Caucasian 32, 153, 155–161, 164, 166–7, 205
- Celtiberian (Cib.) 4, 352, 357, 359–60, 362, 364,  
 367–8, 377
- Celtic (Celt.) 3, 4, 321, 349–50, 353–4, 389,  
 454, 456, 529, 596; further reading 377–8;  
 morphology (nominal) 367–70; morphology  
 (pronouns) 370–1; origins 354–6; P Celtic  
 357–8, 358, 363; phonology 15, 18–19,  
 21, 361–4; position within IE 356–7; Q  
 Celtic 357–8, 363; subdivisions of 357–61;  
 syntax 377; word formation 376–7; *see also*  
 Brythonic (Brittonic/British); Celtiberian;  
 Celtic verbs; Common Celtic; Continental  
 Celtic; Gaelic (Goidelic/Irish); Insular Celtic;  
 Proto-Celtic

- Celtic verbs: active preterit stem 376; future stem 375–6; inflection 371–2; personal endings 372–3; present stem 374–5; subjunctive stem 375
- Cham dialect (Albanian) 50, 567
- Chechen 159, 160
- Chinese 2, 7
- Chinese *Qiuqi* 452
- Chinese *Yanqi* 452
- Choresmian 266
- Church Slavic (CS) 17; *see also* Late Church Slavic; Old Church Slavic; Slavic
- Chukchi 164
- Chukchi-Kamchatkan 153
- Circassian 156
- Classical Latin (Class.Lat.) 28, 33, 50–1, 322–6, 329–38, 341, 343–4, 346–8; *see also* Latin
- Common Baltic (CBalt.) 483, 486; vowels 489; *see also* Baltic
- Common Iranian 264, 266, 269–71
- Common Slavic (CSl.) 16, 482, 489, 521–2, 530, 533–4; *see also* Slavic
- Continental Celtic 4, 352, 353, 358, 362, 368, 370, 376–7; *see also* Celtic
- Continental West Germanic 407; *see also* Germanic
- Cornish (Corn.) 4, 352, 378; Primitive 361, 364
- Cretan (Cret.) 291
- Crimean Gothic 388; *see also* Gothic
- Croatian (Croat.) 3, 519, 522, 545; *see also* Kajkavian Croatian; Serbo-Croatian; Slavic
- Cuneiform Luwian (CLuw.; Luvian) 3, 173–4, 178–9, 181–3; phonology 16, 21; *see also* Luwian
- Curonian 486, 489
- Cymraeg *see* Welsh
- Cymry 353
- Cypriot (Cypr.) 291, 294, 314
- Czech (Cz.) 2, 3, 5, 519, 522; Doudleby dialect 522; Old 17; phonology 15; syntax 545; *see also* Slavic
- Dacian (Dac.) 5
- Daco-Misian 555
- Daco-Romanian 599; *see also* Romanian
- Daic 155
- Danish (Dan.) 3, 394; Old 406
- Dardanian 555
- Dardic 3
- Dargi 159
- Doric Greek (Dor.) 4, 287, 294, 317; *see also* Greek
- Doudleby dialect 522
- Dravidian 1, 7, 153, 155, 157, 159–60, 165, 167, 208–9, 211
- Dutch 3, 5, 393, 403, 405–7; Modern 408; phonology 19; *see also* Old Low Franconian
- Early Modern Nuremberg 397
- East Baltic (EBalt.) 5, 482, 486, 491, 495, 499–502, 505–9, 511–12, 514; dialects 488–9; vowels 490; *see also* Baltic
- East Germanic 5, 29, 388; *see also* Burgund; Germanic; Gothic; Vandal
- East Greek (EGr.) 287, 306–7; *see also* Attic-Ionic Greek; Arcado-Cypriot Greek; Greek
- East Norse 406; *see also* Old Danish; Old Gutnish; Old Swedish
- East Lithuanian (ELith.) 500; *see also* Lithuanian
- East Prussian Lithuanian 488, 493; *see also* Lithuanian
- East Slavic 519, 522, 524–5, 533, 539, 544–6, 557; *see also* Slavic
- Eastern Armenian 421
- Eastern Middle Iranian 266, 283
- Eastern (Balkan) South Slavic 545
- Egyptian 7
- Empire Luwian 173
- English (Eng.) 2, 5, 7, 394, 403, 558; American 400; aspirated stops 15; compared to Italian 8; compared to Spanish 8; Liverpool 394; phonology 15–19, 21; *see also* British; Modern English; Old English
- Epic Sanskrit 242; *see also* Sanskrit
- Eskimo-Aleut 153
- Estonian (Est.) 1
- Etruscan 317, 321, 324, 329
- Faliscan (Fal.) 4, 317
- Faroese 388, 397–8
- Farsi *see* Persian/Farsi
- Finnish/Finnic (Finn.) 1, 159, 389, 522
- Flemish 397
- French (Fr.) 4, 7, 321, 558; Old 389
- Frisian 5, 387
- Gaelic (Goidelic/Irish) 352, 357–8; Irish 4, 352; phonology 365–7; *see also* Middle Irish; Old Irish
- Galatian 352, 359
- Gallo-Brythonic 358; *see also* Brythonic
- Gatha-Avestan (GAv.) 16, 18, 65, 97, 101; *see also* Avestan
- Gaulish (Gaul.) 3, 4, 13, 352–3, 358–61, 364, 367–9, 371–2, 377–8; phonology 16–19
- Georgian 7, 18, 159, 164
- German (Germ.) 3, 5, 8, 407, 596; High 393; Low 5, 16; Middle High 398, 402; phonology 16; *see also* Modern German; Old High German
- Germanic (Gmc.) 4–5, 357, 359, 364, 389, 403, 405–6, 454–5, 483, 488–9, 522, 529; about 387–90; accent shift 396; consonantal shifts, 18, 20, 392; Germanic foot 398; morphology 389, 398–403; noun classes 401; palatalized velars 21; pre-Proto-20; prosody 396–8; sub-groupings of 388; syntax 405–12; word formation 403–5; *see also* Germanic morphology; Germanic phonology; Germanic syntax; Proto-Germanic; West Germanic
- Germanic morphology: inflectional 398–400; nominal system 400–3
- Germanic phonology 16–21, 394–8; consonants 391–4
- Germanic syntax: clausal 407–8; definites 408–9; rise of V2 406; tense, aspect, and mood 410–12; V2 over time 406–7; word order 405–6



- Gheg dialect 5, 552–4, 560, 567; dialectical variations, verbal 592–3; nominal differences 575–6; stress patterns 572; *see also* Albanian
- Gilyak 153
- Goidelic 4
- Gothic (Goth.) 3, 5, 7, 13, 387–9, 396–7, 401–2, 406–7, 409; consonantal shifts 20; Crimean 388; phonology 15, 17–19, 21; and Tocharian 456
- Great Attic Greek 287; *see also* Attic Greek; Attic-Ionic Greek; Greek
- Greek (Gr.) 2, 3, 4, 6, 287, 321, 389, 421–3, 428–30, 447, 454, 456, 481, 526, 545, 558; accent 296–7; Albanian borrowings 557–8; Arcado-Cypriot (AC) 4, 287, 301, 308; aspirated stops 18–19; Attic (Att.) 291, 294, 296, 297, 298–300, 304, 307, 314, 317; Attic-Ionic (AI) 4, 287, 289, 302, 306–7, 313; Byzantine/Medieval 4; consonantal shifts 20; dialects of 317; East 287, 306–7; Grassman's Law 19; Great Attic 287; influence on Latin 349; Ionic 290, 294, 317; Koiné 4, 287, 291, 297, 299, 311, 313; map 288; Modern 4, 297, 558, 595; mood, complementation and subordination 311; morphology (adjectives and adverbs) 302–4; morphology (nominal) 298–302; morphology (pronouns) 304–5; morphology (verbal) 305–10; prepositions 313; Proto- 292, 301; references to Kelti and Galatae 354; syntax 311–14; Thessalian (Thess.) 287, 294, 297; West 287, 307, 313; word formation 313–14; *see also* Ancient Greek; Greek phonology; Homeric
- Greek phonology 15–21, 291–2, 295–297; diphthongs 293; IE fricative 293; labiovelars 292; laryngeals 292; semivowels 292–3; vowel systems 293–4
- Griko 287
- Hamitic 356
- Hattic 164
- Hebrew 2, 3, 166
- Hellenistic Greek *see* Koiné Greek
- Hieroglyphic Luwian (HLuw.) 3, 21, 173
- High German 393; *see also* German; Old High German
- High Latvian (Selonian and Latgalian) 489, 494; *see also* Latvian
- Hindi/Urdu 1, 4, 217
- Hittite (Hitt.) 3, 13, 171–3, 192, 357; Old 178; phonology 15–17, 19, 21
- Homeric (Hom.; Homer) 63, 96–7, 112, 117, 135, 145, 157, 167, 287, 290, 291–2, 298–9, 301–2, 304–5, 308–10, 313–14
- Hungarian 1, 522
- Hurrian 164, 205
- Hurro-Urartian 205
- Icelandic 397, 401; *see also* Old Icelandic
- Illyrian (Illyr.) 5, 555–6
- Indic *see* Indo-Aryan
- Indo-Aryan (Indic) 3, 160, 264; ablaut alternation 230–2; chronology of 214–17; classification of root morphemes 232–3; dentals 227–8; diphthongs 222, 223–5; further reading 252–5, 252–6; general information 214; genetic classification 217–21; Grassman's Law 19; labials 228; Middle Indic 4; morphology (noun) 233–8; morphology (verb) 238–48; morphophonology 229–33; morphotactics 229–30; New Indic 4; palatals 226–7; phonemes 223–6; phonology 15, 18–19, 21, 221–9; plosives 225–6; retroflex 227; sibilants and h 228–9; sonants with vocal variants 222–3; syntax 249–50; and Tocharian 456–7; typological features 250–2; typological features of later languages 250–2; velars 226; vowels 221, 223–5; *see also* Indo-Iranian; Proto-Indo-Aryan
- Indo-European (IE) 7; further reading 8; and Germanic stops 392; Italic 317–18; Late Western 357; morphology 434; phonology 563–4; and Tocharian languages 453–7; verbs 539; *see also* Proto-Indo-European
- Indo-Iranian (Irr.) 3–4, 6, 269, 389, 428, 456; diachronic typological perspective 208–11; further reading 211; lexicon 207–8; morphology 207; origins 205; phonology 21, 206; *see also* Proto-Indo-Iranian
- Ingush 157
- Ingvaeonic 389
- Inscriptional Burmese 389
- Insular Celtic 4, 352–3, 357–8, 363–4, 367–8, 370, 372, 377; *see also* Celtic
- Ionic Greek (Ion.) 290, 294, 317; *see also* Greek
- Ionic-Attic Greek *see* Attic-Ionic Greek
- Iranian (Iran.) 421, 430, 441, 446–7; adjectives 276; Armenian and 422; Avestan 263; consonantism 268; dental affricates 271; Eastern Middle 266; first palatal series 271; further reading 283–4; Modern 263, 277; morphology 266–72, 274–6; morphophonology 272–4; New 4; Old Iranian 4, 272, 283; Old Persian 263; phonology 18, 19, 457; pronouns 276–7; stems with ablaut variation 275; syntax 282–3; and Tocharian 456–7; verbs 277–81; word-formation 281–2; *see also* Avestan; Indo-Iranian; Middle Iranian; Persian/Farsi; Proto-Iranian
- Irish Gaelic 4, 352; *see also* Gaelic; Middle Irish; Old Irish
- Italian (Ital.) 3, 4, 7, 321, 558; compared to English 8
- Italic (It.) 3, 4, 6, 454, 456; epigraphy 350; within Indo-European 317–18; languages of ancient Italy 317; maps 318–19; phonology 18–21; verb system 336; *see also* Proto-Italic
- Italo-Albanian 554, 558
- Italo-Celtic 319–20, 357
- Japanese 85, 159, 165, 172
- Jatvingian 486
- Kafiri 264
- Kajkavian Croatian 519, 527, 538, 542; *see also* Croatian

- Kartvelian 7, 153, 155–6, 162, 166  
 Kashubian 519  
 Ket 155, 156  
 Khotanese (Khot.) 263, 266, 271, 274, 282, 453, 457  
 Kizzuwatna Luwian 173; *see also* Luwian  
 Koine Greek 4, 287, 291, 297, 299, 311, 313;  
   *see also* Greek  
 Korean 32  
*Kucha* 452  
 Kurdish 1, 4  
 Kurzeme 494  
*Kuši* 452
- Lab dialect 552  
 Laconian 291  
 Lak 163  
 Late Avestan 13, 15–17, 19, 283; *see also* Avestan  
 Late Church Slav (LCS) 523; *see also* Church Slav  
 Latgalian 489  
 Latin (Lat.) 2, 3, 4, 7, 8, 13, 21, 317, 596; accent 328–9; alphabets 324–5; borrowings into Albanian 557, 572, 597; case endings 329–33; conjugation classes 337; consonantal shifts 20; declensions 331–3; diachronic developments 339–40; diachronics 326–8, 334; documentary evidence 323–4; lexicon and word formation 349–50; medieval and later 322; morphology (nominal and pronominal) 329–33; morphology (verbal) 336–43; morphophonology 339; non-finite verb forms 341–2; origins of 320; phonology 13, 15–17, 19, 21, 325–6; references to *Keltoi* and *Galatae* 354; and the Romance languages 321–2; spread of 321; varieties of 322–3; verb system 336; *see also* Classical Latin  
 Latino-Faliscan Italic 4  
 Latvian (Latv.) 5, 13, 482, 488, 495, 498; adjectives 503–4; dialects 489; glottal acute 481; Low 489; Old 65; palatalized velars 21; tonal oppositions in 494; verbs 515n3; vowels 491–2; *see also* Baltic  
 Lepontic (Lep.) 4, 357, 359, 362–4, 368  
 Lesbian (Lesb.) 287, 294, 297, 302  
 Letts 489  
 Lezgian 159  
 Ligurian 5  
 Lithuanian (Lith.) 3, 5, 13, 82, 86, 94, 102, 139, 160–1, 480–2, 486–8, 495–6, 498, 526; accent 55, 494; adjectives 503–4; diphthongs 496; nominal stem types 500–1; phonology 15–19, 21, 28–9, 33, 35–6, 38, 41, 46, 48–50, 498; tonal opposition in 492–3; vowels 490–1; Žemaitic dialect 481, 488–9, 496, 508; *see also* Baltic  
 Liverpool English 394; *see also* English  
 Low German 5, 16; *see also* German  
 Low Latvian 489; *see also* Latvian  
 Lower Sorbian (LSorb.) 519  
 Lusitanian 3, 5, 21  
 Luwian (Luw.) 3, 13, 21, 22, 25–7, 171–88, 190, 194  
 Lycian (Lyc.) 3, 21, 22, 26, 27, 171–2, 174–82, 185–8, 190–2  
 Lydian (Lyd.) 3, 171–2, 175, 177–82, 186, 188, 190–2  
 Macedonian (Mac.) 5, 288, 519, 522, 544–7; dialects 557; *see also* Slavic  
 Magadhi 4  
 Maharashtri 4  
 Malayalam 1  
 Maltese 1  
 Mandarin 155  
 Manichean Middle Persian (ManMPers) *see* Middle Persian; Persian/Farsi  
 Manx 4  
 Marathi 4  
 Median 264  
 Messapian/Messapic (Messap.) 5, 321, 357, 555–6  
 Middle Armenian 421; *see also* Armenian  
 Middle Breton (MBret.) *see* Breton  
 Middle Dutch (MDutch) *see* Dutch  
 Middle High German (MHG) 398, 402; *see also* German; Old High German  
 Middle Indic/Prakrits 4; *see also* Indo-Aryan  
 Middle Indo-Aryan (MIA) 250–1  
 Middle Iranian (MIran.) 4, 263, 266, 272, 277, 281; *see also* Iranian  
 Middle Irish (Mir.) 367; phonology 16, 17, 19; *see also* Irish Gaelic  
 Middle Parthian (MParth.) 422; *see also* Parthian  
 Middle Persian (MPers.) 266, 422; *see also* Persian/Farsi  
 Middle Welsh (MW) 15, 17, 352–3, 371, 372, 376; *see also* Welsh  
 Milyan (Lycian B) 3  
 Modern Albanian 553, 558, 560; *see also* Albanian  
 Modern Dutch 408; *see also* Dutch  
 Modern English 399, 403–5; *see also* English  
 Modern German 399, 404, 407–8, 410; *see also* German  
 Modern Greek 4, 297, 558, 595; *see also* Greek  
 Modern Icelandic 401; *see also* Icelandic  
 Modern Iranian 263, 277; *see also* Iranian  
 Modern Standard Albanian 553; *see also* Albanian; Modern Albanian  
 Moldovan 546  
 Mongolic 7  
 Montenegrin 519, 545; *see also* Slavic  
 Munda 160, 208  
 Myanmar 155  
 Mycenaean (Myc.) 29, 291–2, 294, 298, 301–2, 304, 306, 308, 313–14
- Nakh-Daghestanian 205; *see also* Caucasian  
 Navajo 18  
 Nenets 159  
 Nepali 1, 217  
 Netherlandic 396; *see also* Dutch  
 New Indic 4; *see also* Indo-Aryan  
 New Iranian 4; *see also* Iranian  
 New Persian/Farsi 4, 266, 267; nominal inflection 274; *see also* Persian/Farsi

- New Phrygian (NPhryg.) *see* Phrygian  
 Nivkh 156  
 Non-Slavic Balkan 43  
 North Germanic 388–9, 396, 399, 405  
 North Sea Germanic 389  
 North-West Block language 5  
 North-West Greek (NWGr.) 287  
 Norwegian (Norw.) 398; Old 397, 406  
 Nostratic language family 153  
 Nūristānī languages 208, 264
- Old Albanian period 558, 560  
 Old Avestan (OAv. = Gatha-Avestan) 13, 249, 265, 277, 283–4; *see also* Old Iranian  
 Old Bavarian 394  
 Old Breton (OBret.) 354, 361; *see also* Breton; Primitive Breton  
 Old Chinese 2  
 Old Church Slavic (OCS) 5, 7, 13, 53, 66, 82, 86, 95–6, 102, 104, 482, 519, 521, 547; dative absolute 166; phonology 15–19, 21; pronouns 538; syntax 543, 544; *see also* Slavic  
 Old Czech (OCz.) 17; *see also* Czech  
 Old Danish 406; *see also* Danish  
 Old English (OEng.) 5, 13, 387, 396, 401–2, 405–7; consonantal shifts 20; high vowel deletion 398; phonology 16–17, 19  
 Old Frankish 389  
 Old French 389; *see also* French  
 Old Frisian 5, 387  
 Old Gutnish 406  
 Old High German (OHG) 5, 13, 387–9, 394, 396–8, 402, 404, 406–9, 411; phonology 19; *see also* German  
 Old Hittite (OHitt.) 178; *see also* Hittite  
 Old Icelandic (Olc) 16, 388, 406; *see also* Icelandic  
 Old Indic/Indo-Aryan/Indian (OIA) 4, 13, 240, 250–4, 263, 266–7, 268–9, 277, 282–3; phonology 21; *see also* Sanskrit; Vedic  
 Old Iranian 4, 272, 283; *see also* Iranian; Old Avestan  
 Old Irish (OIr) 7, 13, 352–3, 365, 367–71, 376–8, 410; inflectional system 371–2; phonology 15–19, 21; *see also* Irish Gaelic  
 Old Latin (OLat.) 49–50, 118, 322, 325–6, 330, 334–5, 337, 339; *see also* Classical Latin; Latin  
 Old Latvian (OLatv.) 65; *see also* Latvian  
 Old Lithuanian (OLith.) 500, 514–15; *see also* Lithuanian  
 Old Low Franconian (Old Dutch) 5; *see also* Dutch  
 Old Norse (ONor.) 5, 357, 387, 388, 396, 401–2, 404, 407–9; phonology 17–18  
 Old Norwegian 397, 406; *see also* Norwegian  
 Old Persian (OPers.) 3, 4, 13, 263, 264, 266–7, 269, 271, 281, 283–4, 421; morphophonology 272–4; noun cases; pronouns 277; *see also* Persian/Farsi  
 Old Prussian (OPruss.) 5, 13, 456, 481–8, 498–9, 501, 515; tonal oppositions in 495; phonology 15, 19; *see also* Baltic  
 Old Russian (ORuss.) 17; *see also* Russian  
 Old Saxon/Saxonian (OS) 3, 5, 387, 401–2  
 Old Swedish 406; *see also* Swedish  
 Old Welsh (OW) 354, 361, 370; *see also* Welsh  
 Old Welsh, Old Cornish and Old Breton (OWCB) 353, 361, 364–5  
 Oriya 217  
 Oscan (Osc.) 4  
 Osco-Umbrian Italic (Sabellian) 4, 24, 25, 357; *see also* Sabellian languages  
 Ossetic 266
- Palaic (Pal.) 3, 27, 171–3, 178–82, 185–6, 188, 190  
 Pāli 4, 36, 215, 253, 255  
 Pamphylian (Pamph.) 33, 84, 175, 288  
 Panjabi 217  
 Pannonian Slavic 519; *see also* Slavic  
 Parthian 266, 273, 422, 427; Middle 422  
 Pashto 1, 4, 208, 266, 273  
 Pehlevi 4  
 Peloponnesian 287  
 Persian/Farsi (Pers.) 1, 3, 266, 410; Albanian borrowings 558; Middle 267, 422; New 4; *see also* Iranian; Old Persian/Farsi  
 Phrygian (Phryg.) 3, 5, 6, 171, 422, 429, 558  
 Pisdian 3, 175  
 Polabian 519  
 Polish (Pol.) 5, 512, 519, 547; syntax 545; *see also* Slavic  
 Pomeranian 519  
 Pontic 287  
 Portuguese 4, 321  
 Prākritis 4, 25, 215–17, 251  
 Prekmurje Slovene 519, 538, 544; *see also* Slovene  
 pre-Proto-Germanic 20; *see also* Proto-Germanic  
 Primitive Breton 361, 364; *see also* Breton  
 Primitive Cornish 361, 364; *see also* Cornish  
 Primitive Irish 359, 361, 366; *see also* Irish Gaelic  
 Primitive Welsh 361, 364; *see also* Welsh  
 Proto-Albanian 557–60  
 Proto-Anatolian (PAnat.) 49, 171, 190; accent 177; compared to Proto-Indo-European 194; consonants 176–7; lexicon 193–4; vowels 176  
 Proto-Attic-Ionic 294–6  
 Proto-Baltic (PBalt.) 482–3, 496, 521; consonants 496–9  
 Proto-Balto-Slavic (PBSl) 22, 521  
 Proto-Celtic (PCelt.) 18, 24–5, 33–6, 40, 41, 47–8, 51, 171  
 Proto-Germanic (PGmc) 20, 24–5, 30–1, 38, 46, 49–50, 53, 55, 72–4, 79, 85, 171, 388–9, 391  
 Proto-Greek 24–5, 291–4, 299–301  
 Proto-Indian 264  
 Proto-Indo-Aryan 208, 214, 457  
 Proto-Indo-European (PIE): and language typology 153–68; orthography 14; *see also* Indo-European (IE); *see also* Proto-Indo-European morphology; Proto-Indo-European phonology; Proto-Indo-European syntax  
 Proto-Indo-European laryngeals 29, 30, 46–7; \*#HC- (initial pre-consonantal laryngeal) 47–8; \*Ch<sub>1</sub>C 46; \*Ch<sub>2</sub>C 46–7; \*Ch<sub>3</sub>C 47; \*eh<sub>1</sub> 45; \*eh<sub>2</sub> 45–6; \*eh<sub>3</sub> 46; \*h<sub>1</sub>C- 47; \*h<sub>1</sub>e 44; \*h<sub>2</sub>C- 47–8; \*h<sub>2</sub>e 44; \*h<sub>3</sub>C- 48; \*h<sub>3</sub>e 44–5;

- \*HV (laryngeal + vowel) 43–4; influence on neighboring vowels 43; laryngeal metathesis 52–3; resonants and 36–9; syllabic 52
- Proto-Indo-European morphology 158–9, 163; ablaut 61–2; adjectives 79–80; adverbs 80–1; conjunctions 105; further reading 106–7; interjections 106; numerals 89–91; particles 105–6; postpositions 104–5; *see also* Proto-Indo-European nouns; Proto-Indo-European pronouns; Proto-Indo-European verbs
- Proto-Indo-European nouns 78–9; athematic case endings 69–70; athematic stems 62–3, 67–77; athematic stems (*i*-stems) 72–3; athematic stems (*M*-, *l*-, laryngeal, and plosive) 77; athematic stems (*n*-stems) 76–7; athematic stems (*r*-stems) 74–6; athematic stems (*u*-stems) 73–4; case endings 63; case functions 139–143; case syncretism 143–4; case system 158–62; compensatory lengthening in athematic stems 69–70; declension suffixes 62–4; genders of 62, 159–60, 163; heteroclititic stems 77–8; *i*-stems 72–3; *M*-, *l*-, laryngeal, and plosive stems 77; *n*-stems 76–7; *r*-stems 74–6; neuter 63; root nouns 70–2; *s*-stems 78–9; thematic stems 62–7; thematic stems (*Eh*<sub>2</sub>-stems) 62, 66–7; thematic stems (*o*-stems) 62–6; *u*-stems 73–4
- Proto-Indo-European phonology 56, 318; accentual paradigms 526–7; allophones 13; clusters 52–4; diphthongs 564; fricatives 29–30; further reading 56–7; phonemes 13; reconstructed system 13; rules 54–5; stops 13–14, 28, 153, 171, 176; three velar series in 26–7; Tocharian and 454–7; vowels 40–3; *see also* Proto-Indo-European laryngeals; Proto-Indo-European resonants and glides
- Proto-Indo-European pronouns: demonstrative 85–8; indefinite 88; interrogative 88; personal 81–4; possessive 84–5; pronominal adjectives 89; reflexive 84; relative 88–9
- Proto-Indo-European resonants and glides 30–1; asyllabic 31–4; syllabic 34–6; syllabic and laryngeals 36–9
- Proto-Indo-European syntax 111–15; absolute constructions 138–9; argument structure 144–5; case functions 139–44; discontinuous constituents 116–20; further reading 146; positions of the verb 123–38; Wackernagel's Law 120–3; word order 115–16
- Proto-Indo-European verbs 91–2, 468; absolute constructions 138–9; aorist 95–7; imperative 103–4; middle (mediopassive) 99–100; non-finite verb forms 104; optative 102; perfect 97–9; present 92–5; reconstruction of 91–2; subjunctive 100–1
- Proto-Indo-Iranian (PIIr) 205–8, 210–11, 223, 229, 253–4
- Proto-Iranian 208, 214, 264, 269, 457
- Proto-Italic 19, 25, 328–9, 335–6, 349
- Proto-Sabellic 328
- Proto-Semitic 57
- Proto-Slavic (PSl.) 16, 17, 55, 456, 483, 521, 523, 527, 547
- Proto-Tocharian (PToch.) 454, 456–7, 460, 463, 465
- Punjabi 1, 4, 214, 252
- Qarashahr* 452
- Qiuci* 452
- Quechua 18
- Raetic 321
- Romance (Rom.) 317, 389, 393, 454
- Romanian 4, 160, 545–6, 556–7, 559, 579–80, 599–600
- Rumanian 321, 600
- Runic 388–9, 403, 405–7
- Russian (Russ.) 3, 5, 7, 396, 410, 519, 522, 524, 558; syntax 545; *see also* Old Russian; Slavic
- Russian Church Slavic (RussCS) 483, 497, 516; *see also* Old Church Slavic
- Rusyn 519
- Sabellian (Osco-Umbrian) Italic 4, 24, 25, 357
- Sabellic languages 13, 317, 321, 329, 335, 341–3, 349; *see also* Proto-Sabellic
- Saka (Scythian) languages 3, 264, 271, 453
- Samogitian *see* Zemaite dialect
- Samoyedic 159
- Sanskrit (Skr.) 3, 4, 13, 241–2, 263, 393, 458, 470–1, 481; *Agni* 452; Epic 242; grammar 253; and Tocharian 457; Vedic 241–2; voiced stops 17; *see also* Old Indic
- Scandinavian languages 5, 389, 396; *see also* Danish; Finnish; Norwegian; Swedish
- Scots 8
- Scottish Gaelic 4, 352; *see also* Gaelic
- Scythian (Saka) 3, 264, 271, 453
- Selonian 486, 489
- Semitic 7, 57, 90, 157, 159, 162, 166, 356
- Serbian 5, 55, 519, 545; dialects 533, 558; *see also* Slavic
- Serbo-Croatian (SCr.) 5, 481, 527; dialects 557; phonology 16–17; *see also* Bosnian/Croatian/Montenegrin/Serbian
- Shauraseni 4
- Shughni group 266
- Sidetic 3, 171, 175
- Silesian 519
- Sinhala 210, 217, 250, 251
- Sino-Tibetan 153, 155, 159–60
- Slavic (Slav.) 5, 269, 393, 454, 481–2, 486, 489, 503, 514–15, 558; about 519–20; Albanian borrowings 557, 597; deverbatives from CSI 543; disintegration of 522; Eastern (Balkan) South 545; history and migrations 521–2; map 520; metatony 482; micro-languages 519; Pannonian 519; periodization 521; South 524; syntax 543–6; word formation 542–3; *see also* Balto-Slavic; Common Slavic; East Slavic; Proto-Slavic; Slavic morphology; Slavic phonology; Slavic verbs; Western Slavic; Western South Slavic
- Slavic morphology: adjectives 537; adverbs 539; consonantal stem endings 536; indeclinables 539; nominal declension 535–6; nouns 533–4;

- pronouns 537–9; vocalic stem endings 536;  
 vocalic stems 534–5  
 Slavic phonology: accentual paradigms 526–7;  
 cluster simplification 529; consonants 528–33;  
 contraction 525; loss of final consonants 529;  
 loss of post-consonantal j ('deiotation') 531–3;  
 monophthongization of diphthongs 523–4;  
 prosthesis 530; suprasegmentals 525–8;  
 syllabic synharmony and palatalization 524;  
 velar palatalization 530–2; vowels 522–5;  
 Winter's Law 527–8  
 Slavic verbs: ablaut, tense, and aspect 541–2;  
 and Indo-European verb morphology 539;  
 invictions of non-past finite verb forms 540;  
 verb form classes 540–1  
 Slovak (Slk.) 5, 519, 547; syntax 545; *see also*  
 Slavic  
 Slovene (Sln.) 5, 519, 522, 547; Prekmurje 519,  
 538, 544; syntax 545; *see also* Slavic  
 Slovencian 519  
 Sogdian (Sogd.) 4, 263, 266, 274, 281, 284  
 Sorbian, Upper and Lower 519, 547; *see also*  
 Slavic  
 South Picene 317, 341  
 South Slavic 524; *see also* Slavic; Western South  
 Slavic  
 South-West British 352; *see also* Breton; Cornish  
 Spanish (Sp.) 4, 8, 20, 321cx  
 Standard Albanian 558; *see also* Albanian  
 Standard Average European 558  
 Stokavian (Štok.) 33, 519  
 Sumerian 164  
 Swabian 388  
 Swedish (Sw.) 3, 8; Old 406  
 Syriac 421, 423  
  
 Tajik 1  
 Tamian 489  
 Tamil 1  
 Thak 155  
 Thessalian Greek (Thess.) 287, 294, 297; *see also*  
 Greek  
 Thracian, 5, 21, 359, 555–6  
 Tocharian (Toch.) 3, 5, 6, 20, 287, 357, 431,  
 437, 596; about 452–3; characteristics of  
 453–4; development of PIE vowels in 455;  
 further reading 474–5; map 453; morphology  
 461–9, 461–9; place within Indo-European  
 455–8; phonology 15–19, 21, 454–5, 458–61;  
 syntax 471–4; word formation 470–1; *see*  
*also* Proto-Tocharian; Tocharian A, B, and C;  
 Tocharian morphology  
 Tocharian A 5, 13, 452–3  
 Tocharian B 5, 13, 452–3, 457, 472  
 Tocharian C 453, 457  
 Tocharian morphology: adjectival declension 464;  
 adjectives 462–4; declension of 'one' 467;  
 nominal declension 462–3; pronouns 464–7;  
 verbs 467–70  
 Tosk dialect 5, 552–3, 560, 567; Cham 592;  
 dialectical variations, nominal 575–6;  
 dialectical variations, verbal 592–3; stress  
 patterns 572; *see also* Albanian  
 Tsova-Tush (Batsbi) 164  
 Tumshuqese 266, 453, 547  
 Tungusic 7  
 Turkic 1, 7, 158, 208–10; *see also* Altaic  
 Turkish 1, 165, 452–3; Albanian borrowings 558,  
 572, 597  
  
 Ubykh 155  
 Ukrainian (Ukr.) 5, 519, 524, 545, 547; *see also*  
 Slavic  
 Umbrian 4, 15, 317, 321, 329, 341–2, 349  
 Upper/Lower Sorbian (ULSorb.) 519, 547  
 Uralic languages 7, 153  
 Urdu/Hindi 1, 4  
 Uyghur *Kucha* 452  
 Uyghur *Qarashehr* 452  
  
 Vandal 5  
 Vannetais dialect 364  
 Vedic (Ved.) 4, 13, 263; verbal system 238–9,  
 243–5; phonology 15–21, 221; *see also*  
 Indo-Aryan; Old Indic; Sanskrit; Vedic Sanskrit  
 Vedic Sanskrit 241–2; *see also* Sanskrit; Vedic  
 Venetic (Ven.) 5, 21, 318, 321  
 Vidzeme dialect 489, 494  
 Vietnamese 155  
 Volga Finnic 486  
  
 Welsh (W) 2, 3, 4, 352, 363, 368, 377–8;  
 Primitive 361, 364; *see also* Middle Welsh  
 West Baltic 5, 482, 486; *see also* Baltic  
 West British 352; *see also* Welsh  
 West Germanic 388–9, 394, 396, 399, 402,  
 404–5, 407; *see also* Germanic  
 West Greek (WGr.) 287, 307, 313;  
*see also* Greek  
 West Norse 406; *see also* Old Icelandic; Old  
 Norse; Old Norwegian  
 Western Slavic (W Slav.) 519, 524, 531, 545;  
*see also* Slavic  
 Western Armenian 421; *see also* Armenian  
 Western South Slavic 545; *see also* Slavic; South  
 Slavic  
  
 Yanqi 452  
 Yenisean 155  
 Yiddish 393  
 Young Avestan (YAv. = Late Avestan) 13, 15–17,  
 19, 283; *see also* Avestan  
 Younger Avestan 265  
 Yup'ik 163  
  
 Zeelandic 397  
 Žemaitic dialect 481, 488–9, 496, 508  
 Zemgalian 486, 489

---

# SUBJECT-AUTHOR INDEX

---

- ablative absolute 166, 346–7  
ablative case: Albanian 573, 575, 588, 598, 600; Anatolian 178, 180, 182–3, 191; Armenian 435; Baltic 499, 501; Balto-Slavic 479, 482; Celtic 367–8; Germanic 400; Greek 301–2; Indo-Aryan 233, 250–1; Indo-Iranian 207, 210; Iranian 274; Italic 329–31, 333–6, 342; Latin 143, 156; Proto-Indo-European 62–3, 81–2, 138, 140–3, 158–9; Slavic 534, 542; Tocharian 456, 462–3  
ablaut: acrodynamic declension types 68–9, 71, 75, 76; absolute Albanian 573; alternation of 230–2; amphidynamic declension types 68–9, 72–4; Anatolian 178, 187, 189; Armenian 434, 436–7; athematic declension types 67–9; Baltic 490, 508, 510, 515; Balto-Slavic 486n2; Celtic 369, 372; Germanic 391, 398; Greek 300–1, 303, 305, 306; Indo-Aryan 223, 230–3, 239; Iranian 273–6, 278–81; Italic 332; mesodynamic declension types 69, 72, 73; Proto-Indo-European 42, 55–6, 57n10, 61–3, 67–9, 71–8, 92–4, 96, 98, 102, 112, 155, 163; Slavic 539, 541–2; Tocharian 455, 458  
absolute constructions 138–9, 145–6, 166, 168, 346–7, 373  
absolutives 163–4, 238, 250, 257, 303, 588;  
*see also gerunds*  
Açaryan, Hr. 423, 447  
accent: akrostatic 425; Albanian 595; Anatolian 177; Armenian 424; Baltic 495, 504; Celtic 361, 364, 365; circumflex 297; dynamic 155, 273, 296–7, 328; Germanic 393, 396, 398; Greek 296–7; Indo-Aryan 222–3, 232; Indo-Iranian 206; Italic 328–9; Latin 328–9; PIE 55–6; Proto-Anatolian 177; shift 396, 398, 462; Slavic 482, 525–7, 547; synchronic 425; tonal 155, 273, 296–7, 328  
accusative case: Albanian 573–6, 583, 599–9; Anatolian 177–81, 183, 190–2; Armenian 435, 437–8; Baltic 499–502, 506–8, 513; Celtic 359, 368–72; Germanic 400, 402, 404; Greek 310, 312–13; Indo-Aryan 233, 249–52; Indo-Iranian 207, 210; Italic/Latin 329–33, 335–6, 342–3, 345–7; PIE 31, 62, 115, 136, 140, 141, 144–5, 156, 158–60, 162–4, 166–7; Slavic 535; Tocharian 462–3, 465–6, 474  
Achaemenian period 4, 266  
acute tone 479–81, 486, 496, 502  
Adams, Douglas Q. 8, 453, 456, 474–5  
Adams, J. N. 344, 350  
Adiego Lajara, Ignacio-Javier 171  
adjectives: adnominal 470–4; agreement of 62, 252, 462; Albanian 580–2, 598; Armenian 437–40, 446; Baltic 503–5; Balto-Slavic 480; Celtic 360, 368, 370, 374; comparison of 80; denominal 480; deverbal 580; Germanic 401, 403; Greek 301–4, 314; Indo-Iranian 211; interrogative 465; Iranian 276–7, 281; Italic 329–31, 334; Latin 334; PIE 79–80; possessive 180, 470; pronomial 89, 183; Proto-Indo-European 62, 66–7, 73, 79–80, 89, 111; relative 465; Slavic 537; Tocharian 453, 455, 462–4, 472–4; *see also* participles (verbal adjectives)  
adpositions 143, 156–8, 250; Baltic 514; Proto-Anatolian 190–1  
Adrados, Francisco R. 194  
adverbial clauses 312–13; Latin 346–7  
adverbs: Albanian 555, 580–2, 598; Armenian 434, 440; Baltic 499, 501, 512, 513–14; Greek 293, 297, 302–4; Italic 345; local 191; PIE 80–1, 157; pronomial 89; Proto-Anatolian 190–1; Slavic 538–9; verbal 158; *see also* *converbs*  
Aemilia 317  
Afanasevo culture 456  
affricates 14–15, 21–3, 27, 268, 393, 396, 423, 433, 454, 457, 530–3, 568, 571; dental 269, 271, 498  
Afghanistan 3–4, 263  
Africa 214  
Agathangelos 421  
Agbayani, Brian 123, 192  
agent nouns 180, 314, 332, 349, 427, 432, 441, 445, 446, 468, 471, 480, 503  
Ager Faliscus 317  
Agni, Kingdom of 452  
Agricola, Rodolphus (Roelof Huisman) 2  
Ahlqvist, Anders 372  
Aitzemüller, Rudolf 547  
Albania 5, 552  
Albanian alphabet 559  
Albanian diaspora 552, 554, 558, 567  
allative case: Anatolian 190; Baltic 499; Greek 313; Lycian 180; Old Hittite 178; Tocharian 463  
Allen, W. S. 294, 314, 350  
allophones 13, 14, 29, 31–2, 42, 155, 206, 221, 223, 269, 325, 359, 363–6, 566–7  
alphabets: Albanian 559; Armenian 421–2; Avestan 265–6; Cyrillic 519; derived from Greek 171; epichoric 289; Glagolitic 38, 519;

- Greek 289–90; Latin 324–5, 387; Lugano 359;  
 Ogham 361  
*Altindische Grammatik* (Wackernagel) 254  
 ambigender 573  
 Ambrasas, Vytautas 514  
 Anatolia 171–2, 205; *see also* Asia Minor  
 Andersen, Henning 521, 531, 541  
 Andersen, Ø. 290  
 Anderson, Cormac 367  
 Anderson, P. K. 249  
 Andronovos culture(s) 205, 456  
 ‘(the) animals (pl.) run (sg.)’ rule  
 (τὰ ζῷα τρέχει) 65  
 anlaut 365, 565; absolute 365–7, 565  
 Anthony, David 153  
 Antonsen, Elmer 387, 405  
 aorist 95–7, 162, 185, 279, 307–9, 376; Albanian  
 587–60, 593, 595–6; imperative 443–4;  
 indicative 442–3; reduplicated 97; sigmatic  
 96; Slavic 542; subjunctive 444; thematic  
 96–7, 341  
 apocope 187, 192, 360, 373, 425, 435, 515,  
 565, 576  
 Arbëresh writing tradition 554  
 areal patterning 153, 155, 158, 160, 163–4,  
 167–8, 177, 211, 356, 358, 497, 544, 558,  
 573, 599  
 Aristophanes 291  
 Arkadiev, Peter 515  
 Armenia 5, 263  
 Armenian alphabet 421–2  
 Armenian diaspora 421  
 Armenian Republic 421  
 Armenian sound shift 426  
 Armorica 352  
 Arnold, Matthew 353  
 articles: Albanian 574, 580–4; Celtic 367, 370;  
 definite 304, 370, 396, 403, 408–9, 464, 522,  
 558, 579–80, 582, 596; enclitic postponed 440;  
 Homeric 117; indefinite 403, 409, 414n10,  
 440, 558, 574, 600; postpositive 558, 574, 579,  
 582; prepositional 582, 583; prepositive 580,  
 584, 585  
 Arumaa, P. 547  
 Aryans 264  
 Asia Minor 3, 4, 5, 171–2, 174, 287; *see also*  
 Anatolia  
 aspiration 14, 18, 19–21, 30, 45, 47, 57n8, 229,  
 268, 287, 289, 293, 392–3, 423  
 Assamese ergative 210, 249  
 Atharva-veda 215, 217, 252–3  
 Aukštatic areas 493  
 Australia 214  
 Avar Khaganate 521  
 Avery, Peter 392  
 Avestan orthography 265–6  
 Axel, Katrin 132, 396, 397, 406, 407  
 Azerbaijan 263  
 Bactria-Margiana Archaeological  
 Complex 208  
 Baesecke, Georg 410  
 Bakker, E. J. 314  
 Baldi, Philip 350  
 Balkan peninsula 287  
 Balkan Sprachbund 558, 573, 599, 601  
 Balkanindogermanisch 422  
 Balkanisms 558  
 Balkans 5, 171, 555  
 Ball, Martin J. 377  
 Bammesberger, Alfred 399, 413  
 Bangladesh 214  
 Barber, Elizabeth Wayland 474  
 Barbosa, Maria do Pilar Pereira 132  
 Bardhi, Frang (Franciscus Blanchus) 554  
 Barðdal, Jóhanna 145, 211  
 Barrack, Charles 56  
 Bartholomae, Christian 120, 283  
 Bartholomae’s Law 268  
 baryton 180, 189, 297, 486n8, 525–6  
 Basel Epigram 487  
 Bauer, Anna 128, 164, 192  
 Baum, Daniel 242  
 Bayer, Josef 397  
 Beekes, Robert S. P. 106, 146, 163, 240, 284, 430  
 Behagel, Otto 403  
 Behistūn (Bisutūn) inscription 421  
 Belarus 486  
 Bendahman, Jadwiga 107  
 Benelux 5  
 Bengal 205  
 Bergin’s Law 111, 377  
 Bernik, France 547  
 Bhatta, V. P. 255  
 Bhutan 214  
 Biberauer, Theresa 396  
 Binchy, D. A. 377  
 Birchbark Letters 519, 531, 535, 544, 547  
 Bird, Norman 357  
 Birnbaum, Henrik 483, 547  
 Bloch, Jules 254  
 Blom, Alderik H. 353  
 Bloomfield, Maurice 253  
 Boček, Vít 547  
 Böhlingk, Otto 254  
 Boley, Jacqueline 184  
 Bolognesi, Giancarlo 422  
 Bonnet, Guillaume 557  
 Bopp, Franz 3, 112, 479  
 Boretzky, Norbert 558, 588  
 Bousquette, Joshua 397  
 Boxhorn, Marcus van 2–3  
 Braune, Wilhelm 400, 401–2, 404  
 Brereton, Joel P. 253  
 Britain/British Isles 4, 321, 322–3, 355, 389  
 Brittany 4, 352  
 Brixhe, Claude 175, 288  
 Broderick, George 356  
 Bronze Age 320, 486  
 Brugmann, Karl 26, 112, 479  
 “Brugmann’s grade” 239  
 Brugmann’s Law 231, 267, 273  
 Bubeník, Vít 107, 168, 249, 250–1, 255  
 Buccini, Anthony F. 396  
 Buchanan, Scot George 353  
 Buchholz, Oda 572, 573, 588, 603

- Buck, C. D. 314  
 Buddhist texts 215, 452, 454  
 Bulgaria 5, 552  
 Burrow, Thomas 253  
 Buzuku, Gjon 554
- Çabej, Eqrem 603  
 Caesar (Julius) 323, 343, 387  
 Caland, Willem 126  
 Caland System 79, 107  
 Camaj, Martin 597, 599  
 Campania 317  
 Campanile, Enrico 357, 375  
 Campbell, Lyle 8  
 Canada 421  
 cardinal numerals 89, 90, 157, 440, 466–7, 505, 584–5, 600  
 Cardona, George 107, 255  
 Caria 175  
 Carroll, Ryan 399  
 Carruba, Onofrio 192  
 case endings: Anatolian 178–80; Armenian 435; Baltic 501; Celtic 361, 368; Classical Latin 329–30; Indo-Aryan 210, 229, 233–4, 237–8, 251; Italic 329–30, 334, 343; Old Hittite 178; PIE 62–3, 69–70, 163; Tocharian 462; *see also* declensions  
 case functions: in the AIEls and PIE 139–45, 155; Tocharian 454; *see also* ablative case; accusative case; allative case; dative case; ergative case; genitive case; instrumental case; locative case; nominative case; vocative case  
 case loss 387, 390, 400, 412  
 Catullus 323, 325  
 Caucasus 3, 5, 171, 205  
 causal clauses 312, 346  
 causatives 94–5, 167, 189, 223, 238, 245, 278, 340, 374, 399, 404, 445, 460–1, 468–9, 490, 512, 541, 587, 593; morphological 209, 252; oppositions 248  
 Çabej, Eqrem 560, 597  
 Central Europe 5  
 centum languages 6, 21, 497; labiovelars 23–5; loss of labiality 25–6; map 28; palatalized velars 21–3; reflex of PIE \*kw/kw 24; reflexes found in satem languages 27  
 Chadwick, John 291, 297, 314  
 chain shifts 20, 393, 396, 412  
 Chantraine, P. 314  
 Chatterji, Suniti Kumar 254  
 Cheung, Johnny 283  
 China 5, 263, 452  
 Chinese Empire 452  
 Chinese Turkestan (Xinjiang) 263  
 Chomsky adjunction 116  
 Cicero 323  
 Cipo, Kostaq 603  
 circumflex 297  
 circumflex tone: Balto-Slavic 155, 479–82, 489, 491, 496–7, 501–2, 510; Lithuanian 492–3, 495; Old Prussian 495–6; Saussure's Law 493–4, 515n4; Slavic 525–7  
 Civita Castellana 317  
 Clackson, James 56, 107, 146, 323, 344–5, 350, 422, 431  
 clauses: adverbial 312–13, 346–7; alignment of 111, 162–5, 167; adverbial 346–7; causal 312, 346; conditional 346, 407, 473; dependent 472–3, 511, 601–2; independent 105, 471; purpose 346; relative 111, 165–6, 193, 282, 344, 347–8, 364, 472–3, 558; temporal 346; *see also* subordinate clauses  
 clusters: consonant 431–3; dental 54; thorn 53  
 Codex Argenteus 401, 408, 410–11  
 Codex Marianus 535, 538, 544  
 Codex Neumannianus 487  
 Codex Zographensis 538, 544  
 Coetsem, Frans van 399  
 Collinder, Björn 153  
 Colvin, S. 314  
 Common Slavic Linguistic Atlas 547  
 comparative linguistics 3, 7, 112, 217, 250, 252–3, 350, 353  
 comparatives 80, 303; Albanian 580; Armenian 438, 446; Celtic 370; Greek 80, 293, 303–4; Iranian 276; Italic 334; Slavic 537  
 compensatory lengthening 43, 70, 77, 155, 177, 294–6, 328, 366–7, 389, 412, 430, 441, 525  
 complementation: Albanian 601–2; Greek 311, 313; Latin 345–6  
 complementizers 387, 405, 407–8, 412, 414  
 compounding 403, 470, 543, 597, 599; *see also* compounds  
 compounds: Albanian 599; Armenian 438, 446–7; Baltic 505; Celtic 377; Germanic 403; Greek 302, 314; Indo-Aryan 222, 229, 250, 251, 254; Iranian 281; PIE 79, 90; Italic 332, 350; nominal 470–1; Slavic 543; Tocharian 470–1; *see also* compounding  
 Comrie, Bernard 157, 162, 168, 533, 546  
 conditional clauses 346, 407, 473  
 conditional mood 238, 242, 372, 375, 508, 511  
 conjugations *see* verbal conjugations  
 conjunctions: Albanian 586, 601–2; Anatolian 192–3; coordinate 138; Germanic 407; Greek 312; Italic 346–7; Lithuanian 514; PIE 55, 65, 105, 123, 138; postpositive 123; subordinating 146, 184, 347, 407, 412, 538, 544; temporal 312; Tocharian 467  
 consonant clusters: Albanian 462, 465–6, 468–9; Armenian 426, 428, 431, 433; Celtic 372; Germanic 399; Greek 291; Indo-Aryan 229–30; PIE 52–4, 62; Tocharian 455, 459  
 consonant shifts 18, 20, 392–3, 396  
 consonantism 361, 364; Brythonic 365; Celtic 362–3; Gaelic 366–7  
 consonants: Anatolian 176–7; Armenian 426–34; Baltic 496–9; Celtic 361–7; dental affricates 269, 271, 498; dental clusters 54; dentals 14, 28, 35, 54, 227–8, 291, 342, 454, 528; doubled 361; ejective 18, 154; Germanic 391–4, 397; Indo-Iranian 225–9; intervocalic geminate 364; liquids 428; nasals 428–9, 565; palatalized 361, 367, 374, 460, 491, 524; plosive 225–6; pre-Germanic 391; prothetic 529–30, 537;



- Proto-Anatolian 176–7; Proto-Indo-European 52–5; retroflex 225, 227, 229–30, 263, 270, 423; sibilants 177, 206, 217, 228–9, 269, 427–8; Slavic 528–33; *see also* consonant clusters; consonant shifts; lenition; stops
- Constantine 522
- converbs 158, 211, 238, 250, 257; *see also* gerunds
- Corbett, Greville C. 533, 546
- Corsica 321
- Courtenay, Baudouin de 483n1
- Cowgill, Warren 154, 187, 356, 359, 363, 368, 371, 372, 373, 439
- Cretan Hieroglyphic 291
- Croatia 552
- Croft, William 143, 168
- cuneiform 3, 171–4, 177–9, 181–3, 263, 266–7
- Curta, Florin 521
- Cuzzolin, Pierluigi 350
- Cypro-Minoan script 291
- Cyrillic alphabet 519
- Cyrillo-Methodian tradition 519
- Dacia 321
- Daka, Palok 603
- Dalmatian coast 558
- Daniel of Voskopaja 554
- Darius the Great 421
- dative absolute 166, 511, 544
- dative case: Albanian 558, 573–8, 581, 583, 600–1; Anatolian 178–83, 191–2; Armenian 435, 438–9; Baltic 499, 501–3, 507–8, 511–13; Balto-Slavic 6, 480, 482; Celtic 361, 367–8, 370; Germanic 6, 400–2; Greek 144, 301–2; Indo-Aryan 233, 249, 251; Indo-Iranian 207; Iranian 274, 283; Italic/Latin 322–3, 329–31, 333–6, 342–3; PIE 62, 138–9, 141, 143–5, 147, 158, 166–7; Slavic 544
- Daukša, Mykalojus 488
- deaspiration 19–20, 229, 392
- De Bernardo Stempel, Patrizia 359, 360, 362, 364, 369, 370, 376–7
- Debrunner, Albert 254
- declensions: Albanian 574, 576–80, 583; Baltic 500–7; Armenian 429–30; Avestan 275–6; Celtic 368, 370; consonants 235; Greek 287, 298–300, 302–3; holodynamic ablaut 69, 71, 72, 73, 76, 77; holostatic ablaut 69, 78–9; hysterodynamic ablaut 68–9, 75, 77; Indo-Aryan 233–9, 250; Indo-Iranian 210; Iranian 272, 274, 276–7, 281–2; Italic/Latin 329, 331–4, 342, 350; Proto-Indo-European 62–5, 67–8, 77, 81–2, 85, 86, 88, 90, 93, 104, 159; Slavic 533–7, 539; Tocharian 455, 462, 464, 467; *see also* case endings; nouns
- definite articles 304, 370, 396, 403, 408–9, 464, 522, 558, 579–80, 582, 596
- definiteness 387, 403, 405, 408, 409, 504, 537, 542, 546, 573, 574, 576
- definites, Germanic 408–9
- Degener, Almuth 264
- de Hoz, Javier 355
- deictic particles 167, 171, 277, 304, 306, 335–6, 407–9, 508, 538–9, 583–4, 596
- deictic pronouns 182–3, 464–5
- deiotation 524, 530–3
- Delamarre, Xavier 369, 372, 377, 378
- Delbrück, Berthold 111, 120, 124, 127, 249, 254
- Demiraj, Bardhyl 555, 557, 560, 571, 603
- Demiraj, Shaban 560, 573, 577, 578, 579, 603
- demonstratives: Celtic 370–1; Germanic 407–9; Latin 335; PIE 160, 165; *see also* pronouns, demonstrative
- Denmark 389
- dental affricates 269, 271, 498; *see also* affricates
- dental clusters 54
- dentals 14, 28, 35, 54, 227–8, 291, 342, 454, 528
- depalatalization 270, 271, 531
- dependent clauses 472–3, 511, 601–2
- dependent-markedness 156–7
- Derksen, Rick 481, 483, 486n3, 486n6, 515, 547
- desiderative mood 94, 161, 245, 269–70, 340, 376, 508
- desiderative suffix 307–8
- de Simone, Carlo 354
- Desnickaja, Agnija 557, 560, 577, 578
- De Vogelaer, Gunther 397
- devoicing 20, 22, 25, 229, 272, 392–4, 568, 596
- diachronic elements 34, 153, 158, 208–11, 326–8, 334, 403, 443, 475, 515, 595
- diathesis 146, 238–41, 248, 510
- Dickey, Stephen M. 522, 542
- DiGiovine, Paolo 107, 185
- diglossia 323
- Dik, H. 311
- Dillon, Myles 356
- diphthongization 556, 565
- diphthongs: Armenian 424–5; Avestan 266; Baltic 490; Celtic 364, 366; Germanic 395; Greek 291, 293–4; Indo-Aryan 223–5; Indo-Iranian 263; Iranian 266; Latvian 492; Lithuanian 492–3, 496; long 51, 290; monophthongization 523–4, 531; PIE 48–51; spurious 290
- Doluxanov, Pavel M. 521
- Domi, Mahir 603
- Doric future 307
- Dover, K. J. 311
- Dresher, B. Elan 398
- Dressler, Wolfgang 112, 128
- Drinka, Bridget 107, 164
- Dryer, Matthew S. 165
- Duden, Konrad 403
- Duhoux, Y. 314
- Dumézil, Georges 431
- Dunkel, George 192
- Dybo, Vladimír Antonovič 481, 482, 526, 547
- Dybo's Law 357, 486n7, 526
- Džaukjan, G. B. 431
- Èdel'man, Džoj Isosifovna 283
- East Prussia 488
- Eastern Europe 5
- Edda* 406
- Egetmeyer, M. 314

- Eggers, Hans 400, 401–2, 404  
 Eichner, James 176, 177, 178, 186, 187, 189, 194,  
 359, 429, 437, 440  
 Einarsson, Stefán 401  
 ejectives 18, 154  
 Ekblom, Richard 495  
 Elbing Vocabulary 487, 489, 495, 503  
 Elichmann, Johann 2  
 Elizarenkova, Tatjana Jakovlevna 253, 255  
 Elsie, Robert W. 357  
 Emmerick, Ronald 277  
 Emonds, Joseph Embley 389  
*Enchiridion* (Luther) 488  
 enclitic particles 408, 440  
 enclitics 105, 120, 181, 191–2, 270, 272–3, 275,  
 277, 304–5, 335, 344, 396, 408, 438, 507, 514;  
 atonic 297; pronomial 171, 182–3, 397  
 Endzelin, Jan 479, 482, 488, 494–5, 512,  
 514, 515  
 Endzelin's Law 494–5  
 Enger, Hans-Olav 398  
 England 7  
 epigraphy 323–4, 350  
 ergative case 164, 178, 210–11, 249–50  
 ergative construction 65, 163, 167, 210–11, 252,  
 257n29  
 ergative hypothesis 111, 163–4  
 Erlangen School 56  
 Ernout, Alfred 350  
 Eska, Joseph F. 8, 356, 357, 358, 359, 360, 361,  
 363, 368, 370, 373, 376, 377  
 Etruria 317  
 etymology 192, 194, 218, 378, 595; Armenian  
 423; Slavic 547; Tocharian 454  
*Evangelienbuch* 398, 404, 408  
 Evans, D. Ellis 357, 358, 359, 363, 364, 368,  
 376, 378  
 Evans, (Sir) Arthur 291  
 Eythórsson, Þórhallur 406, 407  
 Eznik 421
- Faarlund, Jan Terje 389, 396  
 Faiss, Klaus 284  
 Falerii Veteres 317, 321  
 Falileyev, Alexander 354, 355, 359, 376  
 Fennell, Trevor G. 515  
 Ferrara, S. 314  
 Ferraresi, Gisella 389, 405  
 Fertig, David 397  
 Fiedler, Wilfried 572, 573, 588, 595, 603  
 Fiji 214  
 First Sound Shift 392; *see also* sound changes  
 Fleuriot, Léon 358  
 Forssman, Bernhard 187, 283, 515  
 Forster, P. 389  
 fortis geminates 388  
 fortis stops 396  
 Fortson, Benjamin W. IV 56, 106, 146  
 Fraenkel, Ernst 515  
 France 4, 359, 421  
 Francis, Eric D. 431  
 Frei, Peter 175  
 Freising Folia 519, 531, 544, 547
- Fremdzeichen 458  
 fricatives: Albanian 568; alveopalatal 498;  
 Anatolian 177; Baltic 496–9; Celtic 354; dental  
 532; IE 293; Germanic 381, 389, 391–3, 396;  
 Greek 291–3; Iranian 264, 268, 271; Italic  
 318, 325–8; non-phonemic 14; palatalization  
 532; Proto-Indo-European 13–15, 19–23, 27,  
 29–30, 33, 53; Slavic 528–30, 532–3  
 Friedrich, Johannes 126, 194  
 Friedrich, Paul 112, 165  
 Fritz, Matthias 107, 146, 159  
 Fuß, Eric 397, 407
- Gamkrelidze, Thomas (Tamaz) V. 8, 18, 56, 154,  
 155, 164, 165, 168, 426, 527–8  
 García Alonso, Juan Luis 376  
 García Ramón 287  
 Garde, Paul 547  
 Garrett, Andrew 171, 184, 192, 193  
 Gast, Volker 403  
 Gätters, Alfrëds 515  
 Gelderen, Elly van 409, 410  
 Geldner, Karl Friedrich 253, 283  
 Gelenius, Sigismund (Zikmund Hrubý z Jeleni) 2  
 Gelsen, Henry 515  
 geminates 364; elimination of 528; fortis 388  
 gemination 294, 397  
 genitive absolute 166, 313  
 genitive case: Albanian 558, 573, 575–7, 580–2,  
 599–600; Anatolian 178–80, 182–3; Armenian  
 438–40, 447; Baltic 499–502, 506, 512–14;  
 Balto-Slavic 479–80, 482; Celtic 357, 360,  
 361, 368–71; Germanic 400, 403–4; Greek  
 143, 301–2, 313; Indo-Aryan 233, 249–51;  
 Indo-Iranian 207, 210–11; Iranian 274, 283;  
 Italic/Latin 323, 329–30, 333–6, 342, 347;  
 Italo-Celtic 319–20; PIE 62, 81, 111, 118, 129,  
 138–9, 141, 143–4, 147, 156, 158–9, 163,  
 165–7; Slavic 522, 534–5, 546; Tocharian  
 462–3, 465–6, 472–4  
 genitives: adnomial 512; partitive 513–14, 522  
 Georgia (Soviet) 263  
 Georgiev, V. 555  
 gerundives 190, 238, 246, 250, 251, 342, 348, 470  
 gerunds 146, 158, 238, 250, 342, 445, 511, 586,  
 588, 592; *see also* absolutes; converbs  
 Ghatage, A.M. 255  
 Gimbutas, Marija 8  
 Girdenis, Aleksas 496  
 Gjinari, Jorgji 573  
 Glagolitic alphabet 38, 519  
 glides: geminate 388; hardening 388, 389;  
 intervocalic 299; palatal 388; strengthening  
 397; syllabic 14, 39, 42, 62  
 glottalized stops 18, 154, 167, 426  
 glottallic theory 18, 56–7, 154, 426  
 Gluhak, Alemko 547  
 Godel, Robert 431, 433, 447  
 Goedegebuure, Petra 178, 182, 192, 193  
 Goeman, Ton 397  
 Goldbach, Maria 405  
 Golston, Chris 123, 192  
 Gonda, Jan 124, 238, 249

- Gotō, Toshifumi 211, 245, 248, 253, 254, 283  
 Göttingen Register of Electronic Texts in Indian Languages (GRETIL) 253  
 Government and Binding model 115–16  
   grabar 421  
   grammars: generative 115; historical 253–4  
 Grassmann, Hermann 255  
 Grassman's Law 19–20, 293–4  
 Greece 4, 5, 287, 552  
 Greek alphabet 289–90  
 Greek uncial script 519  
 Green, D.H. 387  
 Greenberg, Joseph H. 111, 115, 153, 165, 168, 538  
 Greenberg, Marc L. 521, 527  
 Greene, David 356, 358, 367  
 Greppin, John A.C. 446  
 Grierson's division 217  
 Griffith, Ralph T.H. 130, 372  
 Grimm, Jacob 3, 112  
 Grimm's Law 20, 353, 392, 393  
 Grunau, Simon 487  
 Gruppenflexion 252  
 Gumo, Kingdom of 453  
 Gusmani, Roberto 175, 192  
 gutturals 154, 225, 559, 571  
 Guyana 214  
 Gvozdanović, Jadranka 107
- Haarmaan, Harald 557  
 Haegeman, Liliane 122  
 Hajnal, Ivo 115, 166, 176, 180, 187, 189, 190  
 Hale, Mark 115, 121, 122, 123, 193  
 Halle, Morris 396  
 Hallstatt Iron Age culture 354  
 Hamp, Eric P. 352, 358, 359, 362, 375, 376, 389, 436, 558, 603  
 Hanna, Patrizia Noel Aziz 398  
 Harbert, Wayne 403, 407  
 Harðarson, Jón Axel 107  
 Harvey, Anthony 361, 364  
 Haspelmath, Martin 558, 599  
 Haug, D. 290  
 Haugen, Einar 397  
 Havlik's Law 525  
 Head Preference Principle (HPP) 409  
 Heather, Peter 388, 389  
 Hecataeus of Miletus 354  
 Heiric of Auxerre 353  
 Held, Warren H., Jr. 193  
 Hemon, Roparz 378  
 Henning, W.B. 266  
 Herodotus 355, 421  
 heteroclitic stems 77–8, 332, 370  
 heteroclitics 81, 100, 159, 167, 437  
 Hethitologie Portal Mainz 194  
 Hettrich, Heinrich 139–41  
 Hewson, John 107, 168, 251  
 hieroglyphic texts 3, 171, 173, 174, 291  
 High German Consonant Shift 393; *see also* sound changes  
 Hildebrandslied 400, 410  
 Hill, Eugen 404, 405
- Hintze, Almut 284  
 Hinüber, Oskar von 255  
 Hirt, Hermann 481  
 Hirt's Law 481, 525  
 Hittites 421  
 Hock, Hans H. 211, 515, 560  
 Hoffmann, Inge 194, 272  
 Hoffmann, Karl 242, 265, 266, 273, 275, 277, 283  
 Hoffner, Harry A. 185, 192, 193, 194  
 Holland, Gary B. 139, 253  
 holodynamic ablaut declension types 69, 71, 72, 73, 76, 77  
 holostatic ablaut declension types 69, 78–9  
 Holst, Jan Henrik 422, 447  
 Holtzmann's Law 388  
 Holub, Zbyněk 522, 527  
 Homer 135–7, 290, 298, 305, 308, 309, 314  
 Homeric *Kunstsprache* 287, 290  
 Hopper, Paul 154  
 Horace 323  
 Horrocks, G.C. 290, 311, 314, 323, 344–5, 350  
 Howell, Robert B. 396, 401, 403  
 Hrubý, Zikmund (z Jeleni) 2  
 Hubbard, Philip 603  
 Hübschmann, Heinrich 447  
 Huggard, Matyas 193  
 Huisman, Roelof (Agricola, Rodolphus) 2  
 Huld, Martin E. 603  
 Humbach, Helmut 284  
 Hurrian empire 264  
 hyperbaton 116–20, 124, 127  
 hypervoicing 413n1  
 hysterodynamic ablaut declension types 68–9, 75, 77
- Iberian Peninsula 4  
 Idsardi, William 392  
 Illič-Svityč, Vladimir 153, 481, 482, 547  
 Illyria 555  
 Illyricum 321  
 imperative mood 161, 281, 338; Albanian 585–91, 597, 601; Anatolian 184, 186, 188; Armenian 428, 441–3; Baltic 508, 510–11; Germanic 405, 407; Greek 305; Indo-Aryan 238–42, 246–7; Indo-Iranian 207; Iranian 278, 281; Italic/Latin 338, 345–6; PIE 91–2, 97, 102–3, 106, 127–8, 161; Tocharian 460, 468, 471  
*i*-mutation 178, 190  
 indefinite articles 403, 409, 414n10, 440, 558, 574, 600  
 indefiniteness 312  
 independent clauses 105, 471  
 India (Republic of India) 3–4, 205; sociolinguistic situation in 214–15, 217  
 indicative mood: Albanian 586–96, 597, 601–2; Anatolian 184–8; Armenian 441–4; Baltic 508; Celtic 372–5; Germanic 411; Greek 305–7, 309, 311–13; Indo-Aryan 238, 242; Indo-Iranian 207; Iranian 277–9; Italic/Latin 323, 338, 340; PIE 6, 91–2, 100–1, 161–2; Slavic 538; Tocharian 468, 473  
 INDOLOGY 253

- Indo-Uralic hypothesis 7  
 infinitives: Albanian 552, 558, 586, 588, 592, 602; Anatolian 178, 180, 190–1; Armenian 445; Baltic 491, 493, 495, 503, 508–13; Balto-Slavic 486; Greek 305, 309–13; Indo-Aryan 222, 238; Iranian 278, 281; Italic/Latin 341–2, 345–6; PIE 104, 107, 165–6; Slavic 540–1, 544; Tocharian 468–9  
 inflection *see* case endings; declensions; verbal conjugations  
 injunctive mood 91, 97, 107, 113, 114, 130, 207, 238–9, 241–2, 246, 277, 279–80, 282, 428, 441–2  
 Insler, Stanley 249, 284  
 instrumental case: Anatolian 180; Baltic 499, 501–2; Balto-Slavic 479; Celtic 361, 367–8; Germanic 400; Greek 139, 201, 302; Hieroglyphic Luwian 182, 191; Hittite 142–3, 178, 183; Indo-Aryan 233, 249–50; Indo-Iranian 207, 210, 211; Iranian 274; Italic 329–30, 342; Lithuanian 513; Vedic 139–41; PIE 62, 81, 139, 143–4, 158, 192; Tocharian 463  
 interjections 106, 158, 440, 558  
 internationalisms 558  
 intervocalic glide 299  
 intrasyllabic harmony 524  
 Iran 3–4, 263  
 Iraq 263  
 Ireland 4, 355  
 Iron Age 320, 354  
 Irslinger, Britta Sofie 376  
 Isaac, G. R. 376  
 Isidor 406  
 Israel 421  
 Italy 4, 5, 287, 359, 552; ancient 317  
 Ivanov, Vjačeslav Vsevolodovič 8, 18, 56, 154, 164, 165, 168, 483, 527–8  
 Iverson, Gregory K. 392, 393, 394, 396
- Jackson, A. V. Williams 263  
 Jackson, Kenneth Hurlstone 356, 361  
 Jacobs, Neil 413  
 Jahowkyan, G. B. 446  
 Jain, Dhanesh 255  
 Jakobson, Roman 153  
 Jamison, Stephanie W. 249, 253  
 Janda, Richard D. 8  
 Janko, Richard 290  
 Jasanoff, Jay H. 107, 184, 187, 368, 369, 372, 374, 375, 376, 427, 596  
 Jaskula, Krzysztof 367  
 Jazyki mira (Languages of the World) series 255  
 Jeffery, L. H. 289  
 jer fall 525  
 Jespersen's Cycle 398  
 Jokl, Norbert 557, 558  
 Jones, Brandtly 290  
 Jones, William 3  
 Jordán Coléra, Carlos 360  
 Joseph, Brian D. 8, 168, 599  
 Joseph, Lionel S. 359, 362, 369, 375, 376  
 Josephson, Folke 192–3
- Joshi, S. D. 255  
 Jutland 389
- Kabelka, Jonas 515  
 Kallio, Petri 486  
 Kammhuber, Annelies 194  
 Kapović, Mate 107, 526, 547  
 Karaliūnas, Simas 490  
 Karl, Raimund 359, 371, 377  
 Karulis, Konstantīns 515  
 Kathol, A. 397  
 Katz, Joshua T. 107, 370  
 Kavalioti, Theodor 554  
 Kazakhstan 457  
 Kellens, J. 278, 283, 284  
 Kempgen, Sebastian 533, 543, 546  
 Kent, R. G. 284  
 Kerswill, Paul 389  
 Keydana, Götz 116, 121, 122, 123, 133, 138–9, 146  
 Killen, John 314  
 Kimball, Sara 176, 187  
 Kiparsky, Paul 115, 396, 407  
 Klein, Daniel 488  
 Klein, Jared S. 128  
 Klingenschmitt, Gert 422, 429, 433, 434, 435, 440, 443, 444, 447, 573, 577, 579, 598  
 Kloekhorst, Alwin 176, 177, 187, 189, 194, 430  
 Kluge, Friedrich 387, 389, 391  
 Kobayashi, Masato 254  
 Kober, Alice 291, 314  
 Koch, John T. 362  
 Ködderitzsch, Rolf 360, 603  
 Kogan, Anton I. 205  
 koineization 389  
 Koivulehto, Jorma 389  
 Koman culture 555  
 König, Ekkehard 403  
 Kopečný, František 547  
 Koriwun 421  
 Kortlandt, Frederik H. H. 18, 163, 240, 245, 427, 428, 430, 434, 435, 448, 481, 486n6, 489, 526, 527, 528, 547  
 Kosovo 5, 552  
 Kostallari, Androkli 603  
 Krause, Wolfgang 387, 474–5  
 Krisch, Thomas 113, 119, 120, 123, 124, 128, 133–4, 135, 138–9, 144, 255  
 Kristoffersen, Gjert 396  
 Kronasser, Heinz 193  
 Kroonen, Guus 389  
 Kroraina, Kingdom of 457  
 Kroraina Prakrit 453  
 Kucha, Kingdom of 452, 453  
 Kulikov, Leonid Igorevič 144–5, 246–7, 248, 254  
 Kümmel, Martin 211, 242, 248, 254  
 Kurgan hypothesis 6–8  
 Kuroda, Susumu 404  
 Kurschat, Friedrich 488, 493  
 Kuryłowicz, Jerzy 363, 364, 367, 409  
 Kurz, Josef 547  
 Küspert, Klaus-Christian 396  
 Kylstra, A.D. 389

- labialization 14  
 labials 14–15, 23, 38–9, 50, 228, 427, 454,  
 489–90, 532–3; *see also* labiovelars  
 labiovelar glides 388  
 labiovelars 14–15, 20, 23–5, 39, 154, 176–7,  
 206, 291–2, 324–5, 327, 363, 366, 388, 391–3,  
 426–7, 441, 480, 486n2, 571  
 Labov, William 8, 396  
 Lahiri, Aditi 398  
 Lambert, Pierre-Yves 352, 358, 360, 361, 368,  
 372, 376, 378  
 Lamberterie, Charles de 439, 447  
 Landi, Addolarata 557  
 Langslow, David 111, 119, 146  
 language typology: phonological 153–5;  
 Proto-Indo-European 153–68  
 Lanman, C. R. 253  
 Laroche, Emmanuel 174  
 laryngeal hardening 443  
 laryngeal metathesis 52–3  
 laryngeals 14, 29, 30, 154, 155, 177, 292;  
 Indo-Iranian 206; interconsonantal 430;  
 PIE 267, 271; syllabic 14  
 Lass, Roger 403  
 Late Common Slavic period 521  
 La Tène Iron Age culture 354  
 Latin alphabet 387  
 Latium 317, 320  
 Law of Illič-Svityč 486n7  
 law of the penult 328  
 Lebanon 421  
 Ledisma, Jacobus 488  
 Lehfeldt, Werner 515, 526  
 Lehmann, Winfred P. 8, 112, 165, 168, 389,  
 408, 413  
 ‘Leiden School’ 56, 526  
 Lejeune, M. 314, 368, 372  
 lenition 177, 361, 364–7  
 Leskien, August 540, 547  
 Leskien’s Law 493, 507  
 Lewis, Henry 361–4, 377, 378  
 lexicon: Albanian 557; Anatolian 193;  
 Indo-Iranian 207–8; Latin 349–50  
 Lhuyd, Edward 353  
 Lieberman, Erez 399  
 Lindeman, Fredrik Otto 362, 440  
 Lindeman’s Law 54, 89  
 linguistics, historical 3  
 Lippert, Jörg 407  
 Livy 323  
 locative case: Albanian 575, 577, 578, 598;  
 Anatolian 178, 180, 183, 190, 191; Armenian  
 435, 439; Baltic, 499, 501, 502–4, 508, 513;  
 Balto-Slavic 482; Celtic 367; Germanic 400;  
 Greek 297, 313; Indo-Aryan 233, 250–1;  
 Indo-Iranian 207, 210; Italic/Latin 329, 331,  
 342; PIE 62, 81, 138–41, 143–4, 156, 158,  
 165; Slavic 522; Tocharian 462  
 Lockwood, W.B. 412  
 Lubotsky, Alexander M. 211, 242, 255, 284  
 Luke (the Evangelist) 166  
 Lugano alphabet 359  
 Lühr, Rosemarie 116, 121, 444  
 Luiten, Tyler 403  
 Lunt, Horace G. 547  
 Luraghi, Silvia 115, 128, 138, 143, 144  
 Luther, Martin 488  
 Lysistrata 291  
 MacAulay, Donald 377  
 Mačavariani 155  
 Macdonell, A. A. 254  
 Macedonia 5, 552  
 Mailhammer, Robert 389  
 Maldives Islands 205, 214  
 Mallory, James P. 8, 153, 171, 354, 456, 474  
 Malzahn, Melanie 475  
 Mann, Stewart 603  
 Manx 352  
 Marek, Christian 175  
 Marggraf’s Law 462  
 Martinet, André 364  
 Martínez, Javier 283  
 Martirosyan, Hrach K. 427, 428, 429, 447  
 Maru 389  
 Masica, Colin P. 211, 252, 255  
 Matasović, Ranko 107, 159, 160, 164, 168, 357,  
 378, 407, 410, 411  
 Mathiassen, Terje 515  
 Matranga, Luca 554  
 Matras, Yaron 252, 601  
 Matzinger, Joachim 422, 429, 447, 455, 554, 555,  
 557, 560, 567, 571, 573, 577, 578, 579, 580,  
 588, 593, 595, 596, 597, 599, 603  
 Mayrhofer, Manfred 154, 211, 264, 273,  
 283, 284  
 Mažiulis, Vytautas 489, 490, 495, 515  
 Mažvydas, Martynas (Martinus Masvidius) 488  
 McCluskey, Stephen C. 487  
 McCone, Kim Robert 111, 124, 352, 356, 358,  
 359, 360, 361, 362, 366, 367, 370, 371, 372,  
 373, 375, 376, 377, 378  
 McManus, Damian 361, 375  
 McWhorter, John H. 389  
 Medes 264  
 Međimurje Kajkavian 544  
 Meid, Wolfgang 125, 194, 354, 356, 357, 358,  
 359, 360  
 Meier-Brügger, Michael 56, 106, 146, 185  
 Meillet, Antoine 26, 167, 350, 389, 428, 433, 435,  
 446–7, 448, 479, 481  
 Meillet’s Law 481, 482  
 Meiser, Gerhard 350  
 Meißner, T. 314  
 Melanesia 214  
 Melchert, H. Craig 171, 176, 177, 178, 180,  
 182, 185, 187, 188, 189, 190, 191, 192,  
 193, 194  
 Mendéz Dosuna, J. 288  
 Menz, Andrea 398  
 Merrill, Peter T. 547  
 ‘Meshari’ 554  
 mesodynamic ablaut declension types 69, 72, 73  
 Mesopotamia 171  
 Mesrop Maštoc<sup>c</sup> 421  
 Messing, G. M. 300

- metathesis 15–17, 52–4, 157, 299, 301, 432, 524, 554  
metatony 481, 482, 493  
Methodius 522  
Meyer, Gustav 579, 603  
Michael Attaleiates 553  
Middle East 3, 5  
Mihăescu, Haralambie 557  
Miller, D. Gary 112  
Milyan 174  
Mitanni kingdom 205, 264, 457  
Mongolia 452  
Monier-Williams, Monier 254  
monosyllabic lengthening 38, 54–5, 70, 83–4, 105  
Montenegro 5, 552  
Morgenstierne, Georg 205, 264  
morphology 7, 8; Albanian 573–97; Armenian 434–45; Baltic 499–512; Balto-Slavic 479, 480; Celtic 367–76; Germanic 389, 398–403; Greek (adjectives and adverbs) 302–4; Greek (nominal) 298–302; Greek (pronouns) 304–5; Indo-Iranian 207; inflectional 398–400; Iranian 266, 274–81; Latin (nominal and pronominal) 329–33; Latin (verbal) 336–43; nominal 178–80, 233–4, 274–6, 298–302, 400–3; Slavic 533–42; Tocharian 461–9, 461–9; verbal 184–90, 238–9  
morphophonology: Indo-Aryan 229–33; Iranian 272–4; Latin 339; morphotactics 229; PIE 112; Tocharian 458–9  
Morpurgo Davies, Anna 177, 185, 301, 314  
Mortensen, David 389  
Moscow Gospel 421  
Moscow School 526  
Movsēs Xorenac'i 421  
Mühlenbach, Karl/ Mülenbachs, Kārlis 488, 494, 515  
Mullen, Alex 361  
Müller, Nicole 377  
Müller, Stefan 391  
Murray, Robert 397, 398  
Mylius, Klaus 255
- Nagorno Karabagh 421  
Narten, Johanna 254, 265, 284  
Narten-imperfects 596  
nasalization 429, 490, 492, 561, 565  
National Awakening/Renaissance (Albania) 558  
Nau, Nicole 489  
near-constraints 54  
Negau helmet 4, 387  
Nepal 214  
Neri, Sergio 107  
Nesselmann, Georg H. F. 486  
Neu, Erich 194  
Newmark, Leonard 603  
*Nibelungenlied* 398  
Nichols, Johanna 156, 157, 160, 167, 168, 521  
Nielsen, Hans Frede 389, 413  
Niepokuj, Mary 107  
Nollé, Johannes 175
- nominative case: Albanian 573–3, 578, 582–3, 599–600; Anatolian, 178–80, 183, 189–90, 192; Armenian 435, 437–8, 444–6; Baltic 499, 502–4, 511–13; Celtic, 368–70, 374, 376; Germanic 400, 402, 404, 411; Greek 297, 312; Indo-Aryan, 233, 250, 252, 255, 257; Indo-Iranian, 207, 210–11; Italic/Latin 329–38, 340–3, 345; PIE 62, 65, 94, 139, 145, 154, 158–60, 162–4, 167; Slavic; Tocharian 462–8, 470, 472–4; *see also* nouns  
Nooten, Barend A. van 253  
Normans 7  
Normier, Rudolf 431  
Norway 389  
Nostratic hypothesis 168  
nouns: abstract 107, 349, 436, 446, 470, 480, 598; agent 180, 314, 332, 349, 427, 432, 441, 445, 446, 468, 471, 480, 503; Albanian 523–30; Armenian 434–8, 445–7; Baltic 499–503; barytone 180, 189, 297, 486n8, 525–6; Common Celtic 367–70; Germanic 400–3; Greek 298–302; Indo-Aryan 233–8; Iranian 274–6; Italic/Latin 329–34; Proto-Indo-European 62–3, 67–74, 77–9, 139–44, 159–60, 163; Slavic 533–6; Tocharian 462–3; *see also* case endings; declensions; nominative case  
Nova Scotia 352  
Novgorod Birchbark letters 519, 531, 535, 544, 547  
numerals: Albanian 584–5; Armenian 440–1; Baltic 505–6; cardinal 89, 90, 157, 440, 466–7, 505, 584–5, 600; Celtic 352–3, 357; Iranian 276–7; ordinal 90, 352–3, 441, 467, 505–6, 580, 585, 600; PIE 89–90, 107, 156–9; Tocharian 466–7
- Ó hUiginn, Ruairi 364, 370  
Oberlies, Thomas 242, 254, 255  
object ellipsis 135  
obstruents 339, 392–4, 412, 510, 529, 568–71  
Oertel, Hanns 249  
Oettinger, Norbert 178, 180, 184, 187, 189  
Ofteidal, Magne 396  
Ogham alphabet 361  
OHG sound shift 20  
Olander, Thomas 106, 547  
Olsen, Birgit Anette 429, 430, 431, 437, 441, 444, 446, 447  
Oman 263  
O'Neil, Wayne 389  
onomatopoeia 106  
optative mood 102, 161, 281, 309, 468, 596–5  
Oranskaja, Tat'jana Iosifovna 255  
ordinal numerals 90, 352–3, 441, 467, 505–6, 580, 585, 600  
Orel, Vladimir 389, 558, 560, 573, 577, 579, 597, 603  
orthography 14, 172, 265–6, 288, 294–5, 323–4, 361, 489–92  
Oscan 317, 321, 328, 329, 330, 342, 349; verb system 336  
Oshiro, Terumasa 188

- Osthoff's Law 301  
 Otfrið 398, 404, 408  
 Otten, Heinrich 126, 127  
 Ottoman invasion 558  
 Ölberg, Hermann 560  
 Ovid 323  
 Oxford Classical Texts series 350  
 oxytone 180, 189, 297, 526
- Pakerys, Antanas 496  
 Pakistan 3, 214, 263  
 Pal Engjëlli (Paulus Angelus) 554  
 Palaic state 173  
 palatal glides 388  
 palatalizations, velar 14–15, 21–3, 26, 28, 154, 206, 226–7, 269, 293, 458–60, 464, 469, 498, 524, 530–2, 534, 571  
 palatovelars 15, 20, 21, 23, 26–7, 206, 496, 498  
 Pāli 255  
 Pamphylia 175, 288  
 Pandit, Shankar Pāndurang 253  
 Parker, H. N. 288  
 Parry, Milman 290  
 participles (verbal adjectives): in the ablative absolute 346; absolute constructions with 165–7; active 180, 238, 247, 281, 446, 511; adjectivization of 599; Albanian 586–8, 590–2, 597, 601–2; aorist 297, 454, 597; Armenian 445; Baltic 511–12; *c*-stem formations in 302; future 445, 516; future active 342; future passive 246; Germanic 398–9, 401, 405, 407; Gothic 410–12; Greek 302, 305, 309, 310–13; Iranian 278–9; Latin 341–2, 346, 348; medio-passive 190, 461; middle 238, 256n24, 281, 446; middle future 309; passive 445; passive perfect 245; past 184, 190, 279, 283, 398, 405, 445; past active 537; past passive 516; perfect 104, 158, 342; perfect passive 211, 249, 251–2, 339; PIE 37, 47, 92, 104–5, 138–9, 158; present active 104, 158, 234, 334, 342, 500, 511, 537; present middle 104, 158, 342; present passive 482; preterit 461, 468–9; Slavic 544, 546; stative 245, 247; Tocharian 464, 468–9, 471
- Patagonia 352  
 patronymics 314  
 P'awstos Buzandac'i 421  
 Pérez Orozco, Santiago 175  
 Pedersen, Holger 8, 194, 361–4, 378, 427, 428, 448, 577, 578, 579, 580  
 periphrastics 184, 310, 339, 342, 358, 396, 404, 411–12, 595  
 Persians 264  
 Petit, Daniel 479, 483, 499, 503  
 Peyrot, Michaël 475  
 phonemes 154, 293, 294; Indo-Aryan 223–5; Latin 324; plosives 225–6; sonants 225; *see also* diphthongs; velars; vowels  
 phonological rules, PIE 54–5  
 phonology 8; Albanian 560–73; Armenian 423–34; Baltic 489–99; Balto-Slavic 479, 480; Brythonic 364–5; Celtic 361–4; Gaelic 365–7; Germanic 391–8, 394–6; Greek 291–7; Hittite 172; of Indo-Aryan languages 221–9; Indo-Iranian 206; Iranian 264, 266–72, 457; Latin 325–6; Slavic 522–33; Tocharian 453, 458–61; *see also* Proto-Indo-European phonology  
 phonotactics 52–4  
 Pinault, Georges-Jean 475  
 Pindar 130–2  
 Pirart, Eric 284  
 Pisani, Vittore 427, 558  
 Pischel, Richard 255  
 Plautus 344, 349  
 Pleterski, Andrej 521  
 Pliny 387  
 Pānianian grammatical tradition 217  
 plosives 225–6  
*Pluralprofilierung* 403  
 poetic language 115; *see also* prosody  
 Pokorný, Julius 356, 378  
 Pokorný's roots 357  
 Polomé, Edgar C. 357, 389  
 Polybius 555  
 polyglossia 216  
 Popko, Maciej 194  
 portmanteau affix 402  
 Portugal 5  
 Porzig, Walter 558  
 postpositions 55, 63, 66, 81, 104–5, 156, 157, 210, 327, 434–5, 462–3, 499, 514  
 Pott, August 479  
 Powell, T. G. E. 354, 355  
 pre-Anatolians 455  
 pre-Indo-Iranians 456  
 prepositions: ablative 301; Albanian 575, 582–3, 598, 600; Armenian 447; Baltic 513–14; Balto-Slavic 480–1; Celtic 358, 368, 377; Germanic 63, 400, 403–4; Greek 301–2, 313; Latin 323, 343–4, 348; PIE 105, 111, 139, 143–4, 146, 156–7; Tocharian 472; *see also* adpositions; postpositions  
 Presamnite 341  
 preterit: with ablaut 461; Anatolian 161–2, 184, 187; Armenian 441; and the aorist 95, 307; Baltic 480, 509–10, 512; Celtic 359, 361, 371, 372, 374–6; dental 359; Germanic 398–400, 404–5, 411–12; Hittite 91, 185–8, 192; Lithuanian 503; non-causative 460; Old Prussian 482; and the perfect 97; sigmatic 461; Slavic 542; suffixless 376; synthetic 542; Tocharian 454, 460–1, 468–9; *see also* verbal tenses and conjugations  
 preterit participles 461, 468–9  
 pre-Tocharians 457  
*Preussische Chronik* 487  
 preverbs 191; Balto-Slavic 480–1; Celtic 377  
 Prifti, Peter 603  
 Probert, Philomen 193, 296  
 Prokosch, Eduard 401–2, 404, 413  
 Prokosch's Law 397  
 Pronk, Tijmen 547  
 pronouns: Albanian 582–4; anaphoric 171, 181–3, 440; Baltic 506–607; Balto-Slavic 482; Celtic 370–1; clitic 507, 514, 600–1; deictic 182–3, 464–5; demonstrative 85–6,

- 181–3, 276–7, 304, 334–6, 370, 440, 507–8, 582–3; dual 401; enclitic 181, 182–3; first and second person 465–6; Greek 304–5; indefinite 88, 183–4, 439–40; interrogative 88, 439–40, 465, 471, 508, 538, 584; Latin 334–6; personal 81–4, 160, 180–1, 277, 304, 334–5, 370–1, 438–9, 506–7, 537–8, 582–3; PIE 81; possessive 84–5, 183, 439, 583–4; pronomial stems 171; reciprocal 439; reflexive 84, 439, 584; relative 88–9, 183–4, 336, 439–40, 465, 471–2, 584; Slavic 537–9; third person 464, 507
- Prosdocimi, Aldo L. 350, 368
- prosody 123; Baltic 492–9; Germanic 396–8
- Prósper, Blanca María 359, 360
- proterodynamic ablaut declension types 68–9
- Proto-Indo-European phonology 56, 318; accentual paradigms 526–7; allophones 13; clusters 52–4; diphthongs 564; fricatives 29–30; further reading 56–7; laryngeals 29–30, 43–53; phonemes 13; reconstructed system 13; resonants and glides 30–9; rules 54–5; stops 13–14, 28, 153, 171, 176; three velar series in 26–7; Tocharian and 454–7; vowels 40–3
- Proto-Indo-European word order: discontinuous constituents (hyperbaton/tmesis) 116–20; marked order 1 (verb–initial V–1) 127–132; marked order 2 (verb–second and verb–third) 133–7; moods 161; unmarked order (verb–final) 133–8; unmarked order (Celtic) 125; unmarked order (Hittite) 236–7; unmarked order (Mycenean Greek) 124; unmarked order (Vedic) 125–6; verb stems 161–2; Wackernagel’s Law 120–4
- prothesis 530
- Ptolemy (Ptolemeu) 354, 555
- Puhvel, Jaan 194
- purpose clauses 346
- Qāwriḡhul culture 456
- quantitative sociolinguistics 8
- Ramat, Paolo 168
- Ramer, Alexis Manaster 437
- Rask, Rasmus 3
- Rasmussen, Jens 176
- Rastorgueva, Vera Sergeevna 283
- Rau, Jeremy 107
- Raybould, Marilynne 355–6
- reduplication: Albanian 456; Anatolian 189; aorist 97, 113–14, 132, 189; Armenian 445; Germanic 399, 405, 413n3; Greek 305, 310; Indo-Aryan 233, 238–40, 242, 248; Iranian 273, 278–9; Latin 337, 340; Old Irish future 375–6; Proto-Indo-European 17, 24, 31, 62, 83–4, 92, 97–8, 104, 187, 539; Slavic 539, 546
- Reichelt, Hans 282, 283
- Reis, Marga 397
- relative clauses 111, 165–6, 193, 282, 344, 347–8, 364, 472–3, 558
- Renan, Ernest 353
- Renfrew, Colin 354
- Renou, Louis 242, 253, 254, 284
- Repaňšek, Luka 376
- Republic of Cyprus 287
- retroflex 225, 227, 229–30, 263, 270, 423
- Ṛgveda (RV) 4, 31, 105, 113, 114, 118–21, 123, 128–30, 140–1, 144–5, 205; *see also* Vedic texts
- rhotacism 29, 165, 325, 327, 334, 389, 552, 566–7
- rhotacization 397
- Riad, Tomas 388
- Rieken, Elisabeth 178, 180, 182, 188, 192, 194
- Rijksbaron, Albert 314
- Ringe, Don 8, 57, 106, 362, 389, 399, 404, 405, 413
- Rix, Helmut 107, 350, 356
- Rizza, Alfredo 192
- Rö Runestone 405
- Roberge, Paul T. 389
- Roberts, Ian 396
- Robinson, Orrin W. 388, 406, 413
- Rodway, Simon 372, 376
- Roma people 3
- Romani 558
- Romania 5
- root constraints 54
- Roques, Mario Luis 554
- Rosen, Hannah 350
- Roth, Rudolf 253, 254
- ruki* rule 6, 29, 206, 229, 269, 428, 441–2, 457, 497–8, 528–30, 537
- Rumsey, Alan 164
- Ruppel, Antonia 138, 313
- Russell, Paul 355, 361, 364, 371, 375, 376, 377
- Russia 7
- Sadnik, Linda 547
- Salmasius, Claudius (Claude de Saumaise) 3
- Salmons, Joseph C. 154, 168, 389, 392, 393, 394, 396
- sandhi 229, 230
- Sardinia 321
- Sasanian period 4, 265
- Sasse, Hans-Jürgen 597
- Sasseti, Filippo 3
- satem languages 6, 21, 426–7, 497, 521, 571; centum reflexes found in 27; labiovelars 23–5; map 28; palatalized velars 21–3
- Saussure, Ferdinand de 493
- Saussure’s Law 493, 515n4
- Schaefer, Christiane 247
- Schallert, Joseph 521
- Schenker, Alexander M. 521, 546
- Schindler, Jochem 429
- Schleicher, August 479, 483, 488
- Schmalstieg, William R. 486n3, 487, 504, 512, 515
- Schmidt, Gernot 107
- Schmidt, Karl Horst 355, 356, 357, 358, 360, 368, 371, 375–6, 375



- Schmitt, Rüdiger 115, 211, 263, 264, 266, 273, 283, 284, 422, 425, 440, 444, 447  
 Schmitt-Brandt, Robert 603  
 Schrijver, Peter 354, 355, 357, 358, 362, 363, 370, 372  
 Schulte, Michael 387  
 Schumacher, Stefan 357, 372, 375, 376, 377, 560, 567, 571, 573, 577, 578, 588, 595, 596, 597, 599, 603  
 Schumann, Roland 107  
*Schwebeablauf* 62  
 Scotland 4, 352–3, 356  
 Second Consonant Shift 393  
 Second Sound Shift 394  
 SEELRC Network 547  
 segmental phonolog problems 57  
 Seldeslachts, Herman 435  
 Selishchev, Afanansij M. 557  
 semivowels 14, 33, 292–3, 424, 429  
 Serbat, Guy 143  
 Serbia 552  
 Shatskov, Andrej 189  
 shibilants 423, 433  
 Shuteriqi, Dhimitër S. 554  
 sibilants 177, 228–9, 427–8; Indo-Aryan 217; Indo-European 269; Indo-Iranian 206  
 Sicilia 321  
 Sidel'tsev, Andrej V. 192  
 Sieg, Emil 474  
 Sievers' Law 54, 396  
 Sihler, Andrew L. 57, 106, 306, 350, 391, 535  
 Silk Road 5  
 Sims-Williams, Patrick 353, 354, 355–6, 358, 359, 360, 361, 363, 364, 365, 366, 367, 368, 370, 372, 373, 375, 376, 377  
 Sinhala instrumental case 210, 249  
 Sintashta archaeological culture 205  
 Skjærvø, P. O. 283  
 Slovenia 5  
 Smith, Laura Catharine 398  
 Smitherman, Thomas 145  
 Smoczyński, Wojciech 515  
 Smyth, H. W. 300  
 Snoj, Marko 521, 538, 543, 547  
 sociolinguistics, quantitative 8  
 Solta, Georg Renatus 423, 447  
 Somers Wicka, Katerina 398  
*sōtēra* rule 297  
 Souček, Vladimír 126, 127  
 sound changes 20, 388, 392–4, 396, 402, 412, 426, 559  
 South America 214  
 South Asia 214  
 South Montenegro 5  
 South-East Asia 214  
 Southworth, Franklin C. 205, 211, 255  
 Spain 4, 321, 359  
 Speijer (Speyer), Jacob S. 254  
 spirantization 364, 365, 367, 392  
 split-ergative alignment 162–3  
 Sprachbund 483  
 Sri Lanka 214  
 Stalmaszczyk, Piotr 357  
 Stang, Christian S. 479, 481, 482, 483, 486n2, 497, 499, 509, 525, 547  
 Stang's Law 70, 526  
 Starke, Frank 178, 180, 191  
 Stearns, McDonald, Jr. 388  
 Steele, P.M. 314  
 stele of Xanthos 174  
 Stempel, Reinhard 107, 446–7  
 Sternefeld, Wolfgang 132  
 Stifter, David 358, 359, 360, 362, 371, 377  
 Stockwell, Robert P. 405, 407  
 stops 176–7, 426–7, 498–9; aspirated 14–15; dental 498, 499; fortis 396; geminate 388; glottalized 18, 154, 167, 426; palatalized 498; PIE 14–29; Proto-Anatolian 176; velar 14, 177, 325, 391, 480, 496, 498, 533; voiced 14–18, 20–1, 54, 153–4, 176–7, 291, 327, 392, 481, 496, 516n7; voiced aspirated stops 14–15, 18–21, 153–4, 176, 206, 222, 264, 268, 291, 326, 392, 426–7, 432, 454, 528; voiceless 14–16, 20, 26, 54, 153–4, 171, 176–7, 225, 268, 359, 392–3, 396, 431–2, 454  
 Strand, Richard F. 205  
 Stressed Syllable Law 397  
 Stuart-Smith, Jane 326  
 Stüber, Karin 369, 370, 377  
*Studi Etruschi* 350  
 Sturtevant, Edgar H. 194, 314  
 subjunctive mood 100–1, 161, 338, 411, 468; in dependent clauses 473; Greek 309, 311  
 subjunctive stem 375  
 subordinate clauses 193; Albanian 601; Anatolian 193; Celtic 397; Germanic 405–6, 408; Greek 305, 311, 312; Indo-Aryan 223; Italic/Latin 343–6, 348; PIE 105, 139, 166; Slavic 546; Tocharian 468  
 superlatives: Albanian 580–1; Armenian 438; Baltic 514; Celtic 357, 370; Greek 303–4; Iranian 276; Italic 334; Italo-Celtic 319; PIE 80  
 Surinam 214  
 Svane, Gunnar 557  
 Svarlien, Diane Arnson 130, 132  
 Sweden 389  
 Syllable Contact Law 397  
 syllable structure, PIE 52–4  
 syncope patterns 371–2, 425  
 syncretisms 138, 143–5, 158–9, 163, 302, 367, 369, 399–400, 434–5  
 syntax: adverbial clauses 346–7; Albanian 599–602; Armenian 447; Baltic 512–15; Celtic 377; Classical Armenian 447; complementation 345–6; configurational 191–3; constituent order 343–4; Germanic 405–12; Greek 311–14; Indo-Aryan 216; Iranian 282–3; Latin 343–9; morphosyntax 191; passive and ergative constructions 249; relative clauses 347–8; sequence of tenses 348; Slavic 543–6; subordinate clauses 193, 344–5; subordinating structures 348–9  
 synthetic passives 389  
 Syria 263, 421

- Szemerényi, Oswald J.L. 56, 106, 107, 154, 156  
 Szemerényi's Law 70, 76
- Tacitus 323, 387  
 Tajikistan 263  
 Talmy, Leonard 158  
 Tarim Basin 452, 453, 456, 474  
 Teeter's Law 112  
 temporal clauses 346  
 Ternes, Elmar 352, 361, 378  
 Thesaurus Indogermanischer Text und Sprachmaterialien (TITUS) 253  
 Thiersch, Craig L. 406  
 Thomai, Jani 603  
 Thomas, Werner 474–5  
 Thomason, Sarah G. 165  
 Thompson, R. J. E. 302  
 thorn-clusters 53  
 Thurneysen, Rudolf 367, 374, 376, 378  
 Tibet 452  
 Tichy, Eva 56  
 Tischler, Johann 193, 194  
 tmesis 111, 116–20, 157, 290, 313, 377  
 Topalli, Kolec 560, 577, 597, 599  
 toponyms 524, 533, 556  
 Toporov, Vladimir Nikolaevič 483, 486  
 Tovar, Antonio 355, 358, 368  
 Trautmann, Reinhold 479  
 Tribulato, Olga 314  
 triple representation 430  
 Trips, Carola 403, 407  
 Trubačëv, Oleg Nikolaevič 522, 547  
 Trudgill, Peter 389  
 Turkey 3, 5, 263, 287, 421  
 Turkmenistan 263
- Uhlenbeck, Christian 163  
 Uhlich, Jürgen 359, 376  
 Ukraine 7, 552  
 Ulfilas (Wulfila) 387  
 Umbria 317  
 umlaut 387, 395–6, 402, 411, 424, 458, 460–1, 564–6, 571, 573, 577, 579, 594  
 unergatives 144  
 United Kingdom 214  
 United States 214, 421  
 Untermann, Jürgen 350  
 Urartians 421  
 Ussishkin, Adam 398  
 Uyghurs 474  
 Uzbekistan 263, 452
- Vaan, Michiel de 283, 350, 560  
 Vaillant, André 163, 496, 547  
 Vătăşescu, Cătălina 557  
 Vedas 4, 217; *see also* Rgveda  
 Vedic texts 218–21, 252–3, 526; *see also* Rg-Veda  
 velar palatalization 14–15, 21–3, 26, 28, 154, 206, 226–7, 269, 293, 458–60, 464, 469, 498, 524, 530–2, 534, 571  
 velar stops 14, 177, 325, 391, 480, 496, 498, 533  
 velars 226; geminate 396; PIE 269; *see also* labiovelars; palatovelars; velar palatalization; velar stops  
 Velupillai, Viveka 168  
 Venantius Fortunatus 353  
 Vendryes, Joseph 370, 371, 377  
 Vendryes' Restriction 377  
 Venetian district 558  
 Vennemann, Theo 389, 397, 398, 434  
 Ventris, Michael 291, 314  
 verbal aspect: Albanian 585–6; Anatolian 184–5, 189; Armenian 441; Baltic 512; Celtic 371, 377; Germanic 405, 410–12, 414; Greek 305, 309–12; perfective 95; PIE 92, 95, 97, 107n8, 146, 161–2, 168; Slavic 533, 541–2; stative 79, 92, 94, 97–9, 144, 164, 246–7, 256–7, 305, 309–10, 370, 374, 442–5, 509, 541; Tocharian 468; *see also* aorist  
 verbal forms and inflections: non-finite 104, 190, 341–2, 445; stem formation 189–90, 444–5; verb stems 161–2; *see also* infinitives; participles (verbal adjectives)  
 verbal moods: conditional 238, 242, 372, 375, 508, 511; desiderative 94, 161, 245, 269–70, 340, 376, 508; injunctive 91, 97, 107, 113, 114, 130, 207, 238–9, 241–2, 246, 277, 279–80, 282, 428, 441–2; optative 102, 161, 281, 309, 468, 596–7; subjunctive 100–1, 161, 280–1, 309, 311, 338, 411, 468, 473; *see also* imperative mood; indicative mood  
 verbal positions: Germanic 405–7; marked order 1 (verb-initial) 127–33; marked order 2 (verb-second/verb-third) 133–7; unmarked order (verb-final) 124–7  
 verbal tenses and conjugations: Albanian 585–97; Anatolian 184–90; Armenian 441–5; Baltic 486–7, 508–12; Celtic 371–6; Germanic 410–12; Greek 305–10; Indo-Aryan 238–47; Indo-Iranian 207; Iranian 277–81; Latin 336–43; old Indian 240–2; Oscan 336; Proto-Anatolian 184–90; Proto-Indo-European 91–7, 99–104, 138–9; Slavic 539–42; Tocharian 454, 467–9; *see also* ablaut; aorist; preterit  
 verbal voices: active 185–8, 207, 337–8, 467–8; mediopassive (middle voice) 99–100, 188, 207, 209, 308, 338, 372, 389, 444, 467–8; passive 162, 245–7, 339, 373–4; proterokinetic 68, 190  
 verbs: athematic roots 239, 306; auxiliary 162, 184, 283, 546, 585–7, 592; 94, 180, 189, 337–8, 340, 350, 374, 376, 399, 404, 443–6, 464, 468, 470; deponent 99, 308, 326, 338–9, 371–2, 374, 376, 591; ditransitive 144; injunctives 91, 97, 107, 108n12, 113, 114, 130, 207, 238–9, 241–2, 246, 277, 279–80, 282, 428, 441–2, 594; intransitive 144–5, 162–4, 167, 190, 248, 256n24, 283, 307–8, 404, 467–8, 472, 483–4, 510, 539, 588, 592; labile 144–5, 209, 248; periphrastic 184, 310, 339, 342, 358, 396, 404, 411–12, 595; transitive 65, 144–5, 162–4, 167, 190, 240, 247–8, 256–7,

- 281, 287, 307–8, 447, 468, 510, 587–8, 600;  
unaccusative 144, 171
- Vergil 323
- Verkerk, Annemarie 158
- Vermeer, Willem 528, 529, 531, 535, 560
- Verner, Karl 393
- Verner's Law 20, 393, 396
- Verschärfung* 388
- Villar, Francisco 164, 360, 364, 368, 372
- Vine, Brent 350
- Vishva Bandhu 255
- Viredaz, Rémy 434
- Viret, Jérémie 360
- vocalism: Brythonic 364–5; Celtic 362;  
Gaelic 366
- vocative case: Anatolian 178, 183; Armenian  
434; Baltic 499, 502; Celtic 368–9;  
Germanic 400, 402; Indo-Aryan 223, 233;  
Indo-Iranian 207; Italic/Latin 330–1, 333; PIE  
62, 80–1, 105–6, 158–9; Slavic 535; Tocharian  
462–3
- voiced aspirated stops 14–15, 18–21, 176, 206,  
222, 264, 268, 291, 326, 392, 426–7, 432,  
454, 528
- voiced stops 15–18, 20–1, 54, 153–4, 176–7,  
291, 327, 392, 481, 496, 516n7
- voiceless fricatives 15, 264, 326, 392–3
- voiceless stops 15–16, 20, 26, 54, 154, 171,  
176–7, 225, 268, 359, 392–3, 396, 431–2, 454
- Vorlage*: Greek 407; Latin 406, 409
- vowels: Albanian 561–2; anaptyctic 266;  
Anatolian 176; Armenian 423–6; auslaut  
565; Baltic 489–92; Celtic 362, 364, 366–7;  
composition 377; contraction of 425–6;  
doubling 324, 328; dummy 291; East Baltic  
490; epenthetic 53, 266; Germanic 394–6;  
Greek 293–6; Indo-Aryan 221–5; Indo-Iranian  
206–7; Iranian 266–8, 272; Italic/Latin  
325–6, 328; Latvian 491–2; lengthening 154,  
324, 328, 397; Lithuanian 490–1; pre-tonic  
364; prothetic 34, 48, 52, 53, 292, 293, 428,  
429, 430, 432, 440; Proto-Anatolian 176;  
proto-Germanic 395; Proto-Indo-European  
39–43, 154–5, 455; Slavic 522–5; Tocharian  
455; weakening of 425; *see also* diphthongs;  
semivowels
- Voyles, Joseph 56
- Wachter, Rudolf 350
- Wackernagel, Jacob 111, 119, 120, 146, 249,  
254, 405
- Wackernagel clitics and particles 116–17, 123
- Wackernagel's Law 120–3, 146, 192, 311, 371,  
373, 377, 543; in Baltic 514–15
- Wagner, Heinrich 356
- Wales 4
- Wallace, Rex 350
- Watkins, Calvert 107, 111, 113, 119, 120, 165,  
192, 194, 247, 356, 357, 358, 370, 372, 373,  
375, 376, 377
- Weerman, Fred 407
- Weight Law 397
- Weise, Oskar 397
- Weiß, Helmut 396, 397
- Weiss, Michael 350, 357, 375
- Weitenberg, Joseph J. S. 180, 192, 435
- Werba, Chlodwig H. 255
- West, Martin L. 282, 284
- West Macedonia 5
- Whaley, Lindsay J. 168
- Whitney, William Dwight 242, 253, 254, 255
- Wiesinger, Peter 396
- Willis, David W. E. 377
- Windfuhr, Gernot 282, 283
- Winter, Werner 428, 430, 452, 462, 481, 527
- Winter's Law 18, 154, 481, 486n6, 489, 506; in  
Baltic 516n13; in Slavic 527–8
- Witczak, Krzysztof T. 357
- Witzel, Michael 253
- Woodard, Roger 194
- word formation: Albanian 597–9; Armenian  
445–7; Celtic 376–7; Germanic 403–5;  
Greek 313–14; Iranian 281–2; Italic  
349–50; Latin 349–50; Slavic 542–6;  
Tocharian 470–1
- word order 164–5, 191–3, 311; *see also*  
Proto-Indo-European word order
- Wujek, Jakub 488
- Wulfila (Ulfilas) 387
- Xhuvani, Aleksander 597
- Xinjiang province 5, 452
- Yakubovich, Ilya 172, 173, 188
- Ylli, Xhelal 557
- Yoshida, Kazuhiko 186, 187, 188
- Young, Steven 495
- Young, Thomas 3
- Zadar 552
- Zair, Nicholas 357, 362
- Zaliznjak, Andrej Anatol'evič 254, 544, 547
- Zarathuštra 4, 457
- Zeilfelder, Susanne 116
- Žemaitija 492–3
- Zemgale 494
- Zeps, Valdis 487
- Ziegler, Sabine 361
- Zimmer, Stefan 192, 377
- Zinkevičius, Zigmas 502
- Žirmunskij [Schirmunski], Viktor  
Maksimovič 388
- Zograf, Georgij Aleksandrovič 252, 255
- Zoroastrian scriptures 264