

KONVENS 2024

**20th Conference on Natural Language Processing
(KONVENS 2024)**

Proceedings of the Conference

September 10-13, 2024

The KONVENS organizers gratefully acknowledge the support from the Austrian Society for Artificial Intelligence.



©2024 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

Introduction

We are thrilled to present the proceedings of this year's installment of KONVENS (Konferenz zur Verarbeitung natürlicher Sprache) held at the University of Vienna from September 10 to 13, 2024. KONVENS is a conference series on computational linguistics established in 1992 organized under the auspices of the German Society for Computational Linguistics and Language Technology, the Special Interest Group on Computational Linguistics of the German Linguistic Society, the Austrian Society for Artificial Intelligence and SwissText.

This year, we received 57 paper submissions, which were peer-reviewed by three reviewers each. Out of all submissions 39 were accepted (21 long papers, 18 short papers). The work presented at KONVENS 2024 spans various topics, including sentiment analysis, question answering, language model evaluation, and the processing of both contemporary and historical languages. We also see an increasing focus on multilingualism, large language models, and the ethical implications of natural language processing technologies. These contributions highlight the ongoing innovation and the importance of addressing both practical applications and theoretical advancements in computational linguistics. This year's conference also features research that tackles the unique challenges of German language processing, alongside studies that explore cross-linguistic applications. The blend of foundational research and applied studies enriches the dialogue within our community and pushes the boundaries of what is possible in language technologies.

We would like to express our heartfelt thanks to all the authors who submitted their work, and a special thanks to the members of the KONVENS 2024 program committee who dedicated their time and expertise to ensure the quality of the conference proceedings in the review process. Your rigorous reviews and thoughtful feedback have been invaluable in maintaining the high standards of the KONVENS conference series. We are excited for the discussions and collaborations that this conference will spark and hope that you find the proceedings insightful and inspiring.

Sincerely,

Brigitte Krenn (Local Co-Chair)

Dagmar Gromann (Local Co-Chair)

Barbara Heinisch (Local Co-Chair)

Michael Wiegand (Workshop, Tutorials and Shared Task Chair)

Pedro Henrique Luz de Araujo (Proceedings Chair)

Benjamin Roth (Program Co-Chair)

Andreas Baumann (Program Co-Chair)

Organizing Committee

Local Chairs

Brigitte Kren, Austrian Research Institute for Artificial Intelligence
Dagmar Gromann, University of Vienna
Barbara Heinisch, University of Vienna

Workshop, Tutorials and Shared Task Chair

Michael Wiegand, University of Vienna

Program Chairs

Benjamin Roth, University of Vienna
Andreas Baumann, University of Vienna

Proceedings Chair

Pedro Henrique Luz de Araujo, University of Vienna

Program Committee

Matthias Aßenmacher, Ludwig-Maximilians-Universität München
Adrien Barbaresi, Berlin-Brandenburg Academy of Sciences
Marcel Bollmann, Linköping University
Peter Bourgonje, Universität Potsdam
Yiyi Chen, Aalborg University
Neele Falk, University of Stuttgart
Dagmar Gromann, University of Vienna
Andrea Horbach, Universität Hildesheim
Bettina M. J. Kern, Technische Universität Wien
Thomas E. Kolb, Technische Universität Wien
Vasiliki Kougia, University of Vienna
Julia Krasselt, ZHAW - Zürcher Hochschule für Angewandte Wissenschaften
Udo Kruschwitz, University of Regensburg
Ekaterina Lapshinova-Koltunski, Universität Hildesheim
Pedro Henrique Luz De Araujo, University of Vienna
Liad Magen, Technische Universität Wien
Thomas Mandl, Universität Hildesheim
Julia Neidhardt, Technische Universität Wien
Friedrich Neubarth, Austrian Research Institute for Artificial Intelligence
Juri Opitz, Ruprecht-Karls-Universität Heidelberg
Rainer Osswald, Heinrich-Heine Universität Düsseldorf
Marek Placiński, Nicolaus Copernicus University
Barbara Plank, Ludwig-Maximilians-Universität München
Jakob Prange, Universität Augsburg
Michael Pucher, Technische Universität Graz
Gabor Recski, TU Wien Institute for Information Systems Engineering
Ines Rehbein, Universität Mannheim
Michael Roth, University of Stuttgart
Josef Ruppenhofer, Fernuniversität Gesamthochschule Hagen
Tatjana Scheffler, Ruhr-Universität Bochum
Yves Scherrer, University of Oslo
Loris Schoenegger, University of Vienna
Marc Schulder, Universität Hamburg
Sabine Schulte Im Walde, University of Stuttgart
Barbara Schuppler, Technische Universität Graz
Anastasiia Sedova, University of Vienna
Melanie Siegel, Darmstadt University of Applied Sciences
Manfred Stede, Universität Potsdam
Andreas Stephan, University of Vienna
Julia Maria Struß, Fachhochschule Potsdam
Lukas Thoma, University of Vienna
Martin Volk, University of Zurich
Ahmadou Wagne, Technische Universität Wien
Christian Wartena, Hochschule Hannover
Michael Wiegand, University of Vienna
Yuxi Xia, University of Vienna

Seid Muhie Yimam, Universität Hamburg
Torsten Zesch, Fernuniversität in Hagen
Erion Çano, Universität Paderborn

Keynote Talk
**Constructions all the way down: rethinking compositionality
in LLMs**

Leonie Weißweiler

UT Austin

Wed, September 11, 2024, 10:00 – 11:00

Abstract: Why are LLMs still not modelling all aspects of language perfectly? Previous works suggested is their deficits in compositionality, regularly building the meaning of an expression as a function of its parts. But in fact, human language is not compositional in this way. Rather, meaning is combined compositionally using constructions, which are pairings of form and function that vary wildly in shape and scope. This means that to achieve the full creativity and flexibility of human language, LLMs will have to assign meaning to constructions and use this to build the meaning of expressions. I will show that this is still not adequately handled by LLMs, and elaborate why construction-compositionality is one of the last remaining challenges that we must solve on our way to more cognitively plausible language models.

Bio: Leonie Weissweiler is a postdoc at UT Austin Linguistics where she works with Kyle Mahowald on the computational learnability of rare linguistic phenomena. She received her PhD from LMU Munich in July 2024, where she worked with Hinrich Schütze on the contributions of Construction Grammar and Morphology to NLP, and vice versa. Her research now focuses on using language models to discover and test hypotheses in Linguistics, while using insights from Linguistics to point out issues with language models.

Keynote Talk

What does it mean for a language model to exhibit a language understanding ability?

Sebastian Schuster

University College London

Wed, September 11, 2024, 14:00 – 15:00

Abstract: Large language models (LLMs) such as GPTs, Gemini or Llama often provide answers that fulfil user requests, which suggests that the model is at least to a large extent able to infer the user's intent and to generate appropriate responses. However, given the open-ended nature of user requests and model responses, it has been quite challenging to systematically evaluate to what extent models exhibit specific language understanding abilities. In my talk, I will focus on one such ability, namely keeping track of how the states of entities change as a discourse unfolds. I will use this ability as a case study for how different evaluation methods can lead to different conclusions about model abilities, I will discuss challenges in evaluating understanding abilities in LLMs and I will consider some recommendations on how to overcome some of these challenges.

Bio: Sebastian Schuster is currently a lecturer in computational linguistics at University College London, and he will start a WWTF-funded research group at the University of Vienna in mid-2025. Before joining UCL, he was a postdoc at New York University and at Saarland University, after completing his PhD at Stanford University. His research focuses on computational semantics and pragmatics and he builds and evaluates computational models of interpreting language in context. His work has won awards at ACL and he has been a senior area chair and program chair at several *ACL conferences and workshops.

Keynote Talk
**Using Natural Language Processing to Advance Social
Science, Responsibly**

Jana Diesner
TU Munich

Thu, September 12, 2024, 09:30 – 10:30

Abstract: Leveraging natural processing techniques to consider the content of information at scale allows us to discover and re-evaluate theories and patterns of societal behavior. This process requires researchers to make a multitude of decisions that require expertise from multiple fields, including how to sample, represent, and preprocess data, implement algorithms, and validate results. I present findings and lessons learned from using NLP techniques, especially entity disambiguation and relation extraction, to study how and why people collaborate and respond to crises. I discuss sources of biases and strategies for mitigating them.

Bio: Jana Diesner is a Full Professor at the Technical University of Munich, School of Social Science and Technology. There, she leads the Human Centered Computing group. Her interdisciplinary group works on methods from network analysis, natural language processing, machine learning and AI, and integrates them with theories from the social sciences to advance our knowledge about complex societal systems and responsible computing. Before joining TU Munich in 2024, she was a tenured professor at the School of Information Sciences at the University of Illinois Urbana Champaign. Jana earned her Ph.D. at Carnegie Mellon, School of Computer Science.

Table of Contents

<i>Large Language Models as Evaluators for Scientific Synthesis</i> Julia Evans, Jennifer D’Souza and Sören Auer	1
<i>A Crosslingual Approach to Dependency Parsing for Middle High German</i> Cora Haiber	23
<i>Complexity of German Texts Written by Primary School Children</i> Jammila Laâguidi, Dana Neumann, Ronja Laarmann-Quante, Stefanie Dipper and Mihail Chifligarov	32
<i>Exploring Automatic Text Simplification for Lithuanian</i> Justina Mandravickaitė, Egle Rimkiene, Danguolė Kalinauskaitė and Danguolė Kotryna Kaplan	40
<i>Word alignment in Discourse Representation Structure parsing</i> Christian Obereder and Gabor Recski	50
<i>Evaluating and Fine-Tuning Retrieval-Augmented Language Models to Generate Text with Accurate Citations</i> Vinzent Penzkofer and Timo Baumann	57
<i>Discourse Parsing for German with new RST Corpora</i> Sara Shahmohammadi and Manfred Stede	65
<i>Revisiting the Phenomenon of Syntactic Complexity Convergence on German Dialogue Data</i> Yu Wang and Hendrik Buschmeier	75
<i>Estimating Word Concreteness from Contextualized Embeddings</i> Christian Wartena	81
<i>Using GermaNet for the Generation of Crossword Puzzles</i> Claus Zinn, Marie Hinrichs and Erhard Hinrichs	89
<i>Leveraging Cross-Lingual Transfer Learning in Spoken Named Entity Recognition Systems</i> Moncef Benaicha, David Thulke and Mehmet Ali Tuğtekin Turan	98
<i>Exploring Data Acquisition Strategies for the Domain Adaptation of QA Models</i> Maurice Falk, Adrian Ulges and Dirk Krechel	106
<i>CO-Fun: A German Dataset on Company Outsourcing in Fund Prospectuses for Named Entity Recognition and Relation Extraction</i> Neda Foroutan, Markus Schröder and Andreas Dengel	117
<i>GERestaurant: A German Dataset of Annotated Restaurant Reviews for Aspect-Based Sentiment Analysis</i> Nils Constantin Hellwig, Jakob Fehle, Markus Bink and Christian Wolff	123
<i>How to Translate SQuAD to German? A Comparative Study of Answer Span Retrieval Methods for Question Answering Dataset Creation</i> Jens Kaiser and Agnieszka Falenska	134
<i>Few-Shot Prompting for Subject Indexing of German Medical Book Titles</i> Lisa Kluge and Maximilian Kähler	141

<i>Binary indexes for optimising corpus queries</i>	
Peter Ljunglöf, Nicholas Smallbone, Mijo Thoresson and Victor Salomonsson	149
<i>An Improved Method for Class-specific Keyword Extraction: A Case Study in the German Business Registry</i>	
Stephen Meisenbacher, Tim Schopf, Weixin Yan, Patrick Holl and Florian Matthes .	159
<i>Tabular JSON: A Proposal for a Pragmatic Linguistic Data Format</i>	
Adam Roussel	166
<i>Semiautomatic Data Generation for Academic Named Entity Recognition in German Text Corpora</i>	
Pia Schwarz	173
<i>Redundancy Aware Multiple Reference Based Gainwise Evaluation of Extractive Summarization</i>	
Mousumi Akter and Santu Karmaker	182
<i>Fine-grained quotation detection and attribution in German news articles</i>	
Fynn Petersen-Frey and Chris Biemann	196
<i>Decoding 16th-Century Letters: From Topic Models to GPT-Based Keyword Mapping</i>	
Phillip Benjamin Ströbel, Stefan Aderhold and Ramona Roller	209
<i>Analysing Effects of Inducing Gender Bias in Language Models</i>	
Stephanie Gross, Brigitte Krenn, Craig Lincoln and Lena Holzwarth	222
<i>OMoS-QA: A Dataset for Cross-Lingual Extractive Question Answering in a German Migration Context</i>	
Steffen Kleinle, Jakob Prange and Annemarie Friedrich	231
<i>Role-Playing LLMs in Professional Communication Training: The Case of Investigative Interviews with Children</i>	
Don Tuggener, Teresa Schneider, Ariana Huwiler, Tobias Kreienbühl, Simon Hischier, Pius Von Däniken and Susanna Niehaus	249
<i>Features and Detectability of German Texts Generated with Large Language Models</i>	
Verena Irrgang, Veronika Solopova, Steffen Zeiler, Robert M. Nickel and Dorothea Kolossa	264
<i>Lex2Sent: A bagging approach to unsupervised sentiment analysis</i>	
Kai-Robin Lange, Jonas Rieger and Jonas Rieger	281
<i>Discourse-Level Features in Spoken and Written Communication</i>	
Hannah J. Seemann, Sara Shahmohammadi, Manfred Stede and Tatjana Scheffler . .	292
<i>Version Control for Speech Corpora</i>	
Vlad Dumitru, Matthias Boehm, Martin Hagmüller and Barbara Schuppler	303
<i>Querying Repetitions in Spoken Language Corpora</i>	
Elena Frick, Henrike Helmer and Dolores Lemmenmeier-Batinić	309
<i>Exploring Phonetic Features in Language Embeddings for Unseen Language Varieties of Austrian German</i>	
Lorenz Gutscher and Michael Pucher	317

<i>A Multilingual Dataset of Adversarial Attacks to Automatic Content Scoring Systems</i>	
Ronja Laarmann-Quante, Christopher Chandler, Noemi Incirkus, Vitaliia Ruban, Alona Solopov and Luca Steen	326
<i>Towards Improving ASR Outputs of Spontaneous Speech with LLMs</i>	
Karner Manuel, Julian Linke, Mark Kröll, Barbara Schuppler and Bernhard C. Geiger	339
<i>OneLove beyond the field - A few-shot pipeline for topic and sentiment analysis during the FIFA World Cup in Qatar</i>	
Christoph Rauchegger, Sonja Mei Wang and Pieter Delobelle	349
<i>Linguistic and extralinguistic factors in automatic speech recognition of German atypical speech</i>	
Eugenia Rykova and Mathias Walther	358
<i>LLM-based Translation Across 500 Years. The Case for Early New High German</i>	
Martin Volk, Dominic P. Fischer, Patricia Scheurer, Raphael Schwitter and Phillip B. Ströbel	368

Program

Wednesday, September 11, 2024

08:30 - 09:30 *Registration*

09:30 - 10:00 *Opening*

10:00 - 11:00 *Keynote 1, Leonie Weißweiler*

11:00 - 12:00 *Poster Session 1 and Coffee Break*

Large Language Models as Evaluators for Scientific Synthesis

Julia Evans, Jennifer D'Souza and Sören Auer

A Crosslingual Approach to Dependency Parsing for Middle High German

Cora Haiber

Complexity of German Texts Written by Primary School Children

Jammila Laâguidi, Dana Neumann, Ronja Laarmann-Quante, Stefanie Dipper and Mihail Chifligarov

Exploring Automatic Text Simplification for Lithuanian

Justina Mandravickaitė, Egle Rimkiene, Danguolė Kalinauskaitė and Danguolė Kotryna Kapkan

Word alignment in Discourse Representation Structure parsing

Christian Obereder and Gabor Recski

Evaluating and Fine-Tuning Retrieval-Augmented Language Models to Generate Text with Accurate Citations

Vinzent Penzkofer and Timo Baumann

Discourse Parsing for German with new RST Corpora

Sara Shahmohammadi and Manfred Stede

Revisiting the Phenomenon of Syntactic Complexity Convergence on German Dialogue Data

Yu Wang and Hendrik Buschmeier

Wednesday, September 11, 2024 (continued)

Estimating Word Concreteness from Contextualized Embeddings

Christian Wartena

Using GermaNet for the Generation of Crossword Puzzles

Claus Zinn, Marie Hinrichs and Erhard Hinrichs

12:00 - 14:00 *Break*

14:00 - 15:00 *Keynote 2, Sebastian Schuster*

15:00 - 16:00 *Poster Session 2 and Coffee Break*

Leveraging Cross-Lingual Transfer Learning in Spoken Named Entity Recognition Systems

Moncef Benaicha, David Thulke and Mehmet Ali Tuğtekin Turan

Exploring Data Acquisition Strategies for the Domain Adaptation of QA Models

Maurice Falk, Adrian Ulges and Dirk Krechel

CO-Fun: A German Dataset on Company Outsourcing in Fund Prospectuses for Named Entity Recognition and Relation Extraction

Neda Foroutan, Markus Schröder and Andreas Dengel

GERestaurant: A German Dataset of Annotated Restaurant Reviews for Aspect-Based Sentiment Analysis

Nils Constantin Hellwig, Jakob Fehle, Markus Bink and Christian Wolff

How to Translate SQuAD to German? A Comparative Study of Answer Span Retrieval Methods for Question Answering Dataset Creation

Jens Kaiser and Agnieszka Falenska

Few-Shot Prompting for Subject Indexing of German Medical Book Titles

Lisa Kluge and Maximilian Kähler

Binary indexes for optimising corpus queries

Peter Ljunglöf, Nicholas Smallbone, Mijo Thoresson and Victor Salomonsson

Wednesday, September 11, 2024 (continued)

An Improved Method for Class-specific Keyword Extraction: A Case Study in the German Business Registry

Stephen Meisenbacher, Tim Schopf, Weixin Yan, Patrick Holl and Florian Matthes

Tabular JSON: A Proposal for a Pragmatic Linguistic Data Format

Adam Roussel

Semiautomatic Data Generation for Academic Named Entity Recognition in German Text Corpora

Pia Schwarz

16:00 - 17:00 *Oral Session 1*

Redundancy Aware Multiple Reference Based Gainwise Evaluation of Extractive Summarization

Mousumi Akter and Santu Karmaker

Fine-grained quotation detection and attribution in German news articles

Fynn Petersen-Frey and Chris Biemann

Decoding 16th-Century Letters: From Topic Models to GPT-Based Keyword Mapping

Phillip Benjamin Ströbel, Stefan Aderhold and Ramona Roller

17:00 - 19:00 *Social event*

19:00 - 21:00 *Dinner*

Thursday, September 12, 2024

08:30 - 09:30 *Registration*

09:30 - 10:30 *Keynote 3, Jana Diesner*

10:30 - 11:00 *Coffee Break*

11:00 - 12:00 *Oral Session 2*

Analysing Effects of Inducing Gender Bias in Language Models

Stephanie Gross, Brigitte Krenn, Craig Lincoln and Lena Holzwarth

OMoS-QA: A Dataset for Cross-Lingual Extractive Question Answering in a German Migration Context

Steffen Kleinle, Jakob Prange and Annemarie Friedrich

Role-Playing LLMs in Professional Communication Training: The Case of Investigative Interviews with Children

Don Tuggener, Teresa Schneider, Ariana Huwiler, Tobias Kreienbühl, Simon Hischer, Pius Von Däniken and Susanna Niehaus

12:00 - 14:00 *Break*

14:00 - 15:00 *Oral Session 3*

Features and Detectability of German Texts Generated with Large Language Models

Verena Irrgang, Veronika Solopova, Steffen Zeiler, Robert M. Nickel and Dorothea Kolossa

Lex2Sent: A bagging approach to unsupervised sentiment analysis

Kai-Robin Lange, Jonas Rieger and Jonas Rieger

Discourse-Level Features in Spoken and Written Communication

Hannah J. Seemann, Sara Shahmohammadi, Manfred Stede and Tatjana Scheffler

15:00 - 16:00 *Poster Session 3 and Coffee Break*

Thursday, September 12, 2024 (continued)

Version Control for Speech Corpora

Vlad Dumitru, Matthias Boehm, Martin Hagemüller and Barbara Schuppler

Querying Repetitions in Spoken Language Corpora

Elena Frick, Henrike Helmer and Dolores Lemmenmeier-Batinić

Exploring Phonetic Features in Language Embeddings for Unseen Language Varieties of Austrian German

Lorenz Gutscher and Michael Pucher

A Multilingual Dataset of Adversarial Attacks to Automatic Content Scoring Systems

Ronja Laarmann-Quante, Christopher Chandler, Noemi Incirkus, Vitaliia Ruban, Alona Solopov and Luca Steen

Towards Improving ASR Outputs of Spontaneous Speech with LLMs

Karner Manuel, Julian Linke, Mark Kröll, Barbara Schuppler and Bernhard C. Geiger

OneLove beyond the field - A few-shot pipeline for topic and sentiment analysis during the FIFA World Cup in Qatar

Christoph Rauchegger, Sonja Mei Wang and Pieter Delobelle

Linguistic and extralinguistic factors in automatic speech recognition of German atypical speech

Eugenia Rykova and Mathias Walther

LLM-based Translation Across 500 Years. The Case for Early New High German

Martin Volk, Dominic P. Fischer, Patricia Scheurer, Raphael Schwitter and Phillip B. Ströbel

16:00 - 17:00 *GSCL PhD Award*

17:00 - 17:30 *Closing*