

Annette Klosa-Kückelhaus, Ilan Kernerman

Lexicography of Coronavirus-related neologisms: An introduction

1 Background

This volume of *Lexicographica. Series Maior* focuses on lexicographic neology and neological lexicography concerning COVID-19 neologisms, featuring papers originally presented at the third Globalex Workshop on Lexicography and Neology (GWLN 2021¹). GWLN 2021 was held online in conjunction with Australex 2021,² with a focus on neologisms arising in relation to the COVID-19 pandemic. Papers discussing various issues related to the detection of such neologisms – including new words, new meanings of existing words, and new multiword units – and their representation in lexicography and dictionaries were invited to offer cross-world views on lexicographic detection and representation of Coronavirus-driven neologisms for different languages. Similar challenges regarding COVID-19 neologisms and lexicography arise for any contemporary language, for example how to detect such neologisms (corpus analysis and editorial means of identification, evaluation of other data, e.g. blogs and chats) or how dictionary users can help with finding and informing about them. But also the extent of borrowing COVID-19 neologisms from other languages (and which ones), in contrast to the use of word formation processes to enlarge the Coronavirus-related vocabulary in a specific language, needs to be examined, and questions of prescriptive vs. descriptive lexicographic information on such neologisms need to be addressed.

The GWLN series began as a single event conjugated with the 22nd Biennial Meeting of the Dictionary Society of North America (DSNA) at Indiana University, Bloomington, in 2019³ and included thirteen invited papers from around the world, of which eight formed a special issue of the DSNA's journal *Dictionaries*, published the following year (2020, 41.1⁴). GWLN-2⁵ was planned in conjunction with the Euralex 2020 Congress (Alexandropoulos, Greece), but due to the COVID-19 pandemic

1 <https://globalex2021.globalex.link/> (last access: 10 June 2022).

2 <https://www.adelaide.edu.au/australex/> (last access: 10 June 2022).

3 <https://dictionarysofamerica.com/> (last access: 10 June 2022).

4 <https://dictionarysofamerica.com/wp-content/uploads/2020/05/Dictionaries-41.1-Table-of-Contents.pdf> (last access: 10 June 2022).

5 <https://globalex2020.globalex.link/gw-euralex2020> (last access: 10 June 2022).

Annette Klosa-Kückelhaus, Leibniz-Institut für Deutsche Sprache, e-mail: klosa@ids-mannheim.de
Ilan Kernerman, K Dictionaries, e-mail: ilan@kdictionaries.com

it was partially held online (November 2020)⁶ and as a special session at Euralex 2020 online (in 2021),⁷ with selected papers published as a special issue of *International Journal of Lexicography* (Klosa-Kückelhaus/Kernerman 2021).⁸

Lexicography has been around for thousands of years and has always had to adapt to developments in society and language, apparently more than ever in the last generation with its increasingly rapid and radical technological changes. Neology has been there forever, driving language from the start and so-to-speak inciting lexicography. Likewise, in recent decades neology has been drawing more attention in research communities and inspiring new practical applications, such as at university or national language observatories or in the language technology industry, as well as with the general public. The speed of novelty in daily life accelerates and the volume of innovations grows exponentially – all defined by language as well as affected by and affecting language. Altogether, there is greater interest in neologisms and in the role of lexicographic resources to capture and disseminate them to the world.

The overall aim of GWLN and its corresponding publications is to explore this intersection of neology and lexicography worldwide, uncover the common factors and highlight individual features, expose and share the findings with each other and enhance mutual understanding, professional competence, and user satisfaction. The main issues in question begin with the identification of neologisms and go on to comprise their categorization and lexicographic treatment and representation. As such, the description in our introduction to the special issue of *Dictionaries* (Klosa-Kückelhaus/Kernerman 2020) is appropriate here, too, and we reproduce it with slight adjustments:

“Neology constitutes a natural, dynamic and multilateral part of all living human languages, whether as a reflection or for facilitation of linguistic communication, and lexicographic interest in neologisms is at least as old as dictionaries themselves. There is a vast field of research of neologisms, pertaining to their origin (stemming from the given language as in new word formation, or loan words from other languages including the dominance of English today, as well as combining both), distribution (in general language and in domain-specific language, that is terminology), identification (applying corpus linguistics methods, editorial methods, user generated candidates, and comparison of different methods), evaluation (such as in blogs and chats), and more. The general definition of neologisms as applied here refers to new words, new multiword units, new elements of word formation, and new meanings of either of them, and addresses lexicography-driven or -oriented aspects, including:

⁶ For the program, see <https://globalex2020.globalex.link/globalex2020-online/> (last access: 10 June 2022).

⁷ <https://euralex2020.gr/> (last access: 10 June 2022).

⁸ <https://academic.oup.com/ijl/issue/34/3> (last access: 10 June 2022).

- How to interoperate lexicographic datasets with online resources and incorporate neologisms into dictionaries (the media, formatting, labelling, etc.)
- How to deal with grammatical/orthographic/pronunciation variation (descriptive vs. prescriptive approaches)
- How to explain meaning with/without encyclopaedic information, and how to use illustrations and audio-visual media
- How well are neologisms that are integrated in dictionaries accepted by the community (issues of rejection of new words and language purism)
- How differently, if at all, should neologisms be treated in different dictionary types (e.g. in historical comprehensive ones as opposed to those focusing on current usage; in monolingual vs. bilingual dictionaries; in special dictionaries of neologisms; in special domain dictionaries)
- How to deal with neologisms that are no longer *new* and with those no longer used
- How can dictionary users help with finding and informing about neologisms

The papers in this volume pursue the discussion on some of these aspects, presenting state-of-the-art research into neology [specific to the COVID-19 pandemic] and ideas on modern lexicographic treatment of neologisms in various dictionary types.”

2 This volume

The thirteen papers in this volume focus on ten languages: one Altaic (Korean), one Finno-Ugric (Hungarian), two Germanic (English and German), four Romance (French, Italian, [Brazilian and European] Portuguese and [Pan-American and European] Spanish), and one Slavic (Croatian), as well as the Sign Language of New Zealand. Specialized dictionaries of neologisms are discussed as well as general language ones, monolingual, bilingual and multilingual lexical resources, print and electronic dictionaries. Questions regarding terminology as well as general language and standard and norm regarding COVID-19 neologisms are raised and different methods of detecting candidates in media corpora, as well as by user contributions, are discussed.

The papers are broadly arranged in four groups of three (and four) papers each. The first group features papers regarding English, German, and Korean, respectively, evolving from systemic neological and lexicographic research carried out in their authors’ institutions for some years, which conveys solid support and wide perspectives to their findings. The second consists of three papers regarding Spanish neologisms in traditional and upcoming lexicographic contexts from Europe and Latin America. The third presents work on Croatian, Hungarian, Italian, and Portuguese in Portugal and Brazil, i.e. to some extent lesser used languages, which is no less pertinent as for dealing with similar issues. The fourth group of papers extends beyond mainstream lexicography to study COVID-19 neology in relation to

collaborative editing in Wiktionary, to terminology, and to New Zealand sign language. Together, this collection offers rich insights that sometimes overlap while remaining unique.

In *The Oxford English Dictionary and the language of Covid-19*, Danica Salazar and Kate Wild offer insight into how the editorial team working on this renowned historical dictionary of English reacted to the challenges posed by the rapid expansion of new vocabulary during the Coronavirus pandemic: “The lexical adaptation necessitated by this global health crisis has been unprecedented in speed and scope, and in response, the *Oxford English Dictionary* (OED) has continually revised its coverage, publishing special updates of Covid-19-related words in 2020 outside of its usual quarterly publication cycle.” The *Oxford Languages’ monitor corpus of English* and other text databases were used to monitor the development of pandemic-related words, and the authors describe how new lexemes (most prominently *COVID-19*) and words with new meanings (e.g. *bubble*) or new significance (e.g. *social distancing*) were detected and treated lexicographically, for example by revising existing entries and adding new ones. Questions of how the use of terminology in general discourse and regional variation should be transferred into lexicographic information are discussed as well.

Finally, the authors explain how their work expanded beyond the dictionary itself: “The OED’s efforts to document the lexical change brought to the English language by the coronavirus pandemic continued throughout 2020, culminating with the *Words of an Unprecedented Year* report, which was published at the end of the year in place of the usual selection of a single Word of the Year (Oxford Languages 2020). This expansive report on the words that defined 2020 features an entire section dedicated to the language of Covid-19.” Many dictionary projects around the world reacted in a similar manner and started to publish texts on the COVID-19 vocabulary addressed to the public, as other papers in this volume show.

While the OED as a comprehensive dictionary on general language will only include some highly frequent new lexemes or new meanings into its content, specialized dictionaries on neologisms can be more generous when it comes to the number of new entries. In the paper titled **German Corona-related neologisms and their lexicographic representation**, Annette Klosa-Kückelhaus discusses this question and contrasts two different perspectives: “There are some (neologism) dictionaries that only record neologisms retrospectively, that is after their lexicalization. [. . .] Other neologism dictionaries [. . .] record neologisms [. . .] before they are fully lexicalized, but are nevertheless accepted parts of the lexicon.” Presenting data on an online neologism dictionary published by the Leibniz Institute for the German Language (IDS), the author demonstrates how both approaches are combined in one project so that dictionary users may find information on COVID-19 neologisms (new lexemes, new meanings, and new usages) as soon as possible throughout the pandemic development. She also discusses how to detect candidates for inclusion, for example

by continuously evaluating user contributions via a word proposal form on the dictionary webpage.

Overall, this dictionary project seems to have profited from the challenges posed by the rapid vocabulary expansion throughout the pandemic, as “the general awareness for the lexicographic work at IDS and for the usefulness of reliable, up-to-date dictionaries” was raised, “making it worthwhile to immerse in lexicography ‘at the pulse of time’.”

The emergence and spread of Korean COVID-19 neologisms in news articles and user comments and their lexicographic description is the topic of the paper by Kilim Nam, Jinsan An and Hae-Yung Jung, in which they examine the occurrence frequencies and usage trends of COVID-19 neologisms in news articles and user comments related to the pandemic, to provide information on Korean neologism usage across genre. As “COVID-19 neologisms, in particular, have proliferated for the past year or so, to express, describe, and comment on a global phenomenon, constituting an unprecedented case of profuse and multifaceted neological creativity centered on a single topic”, they lend themselves especially well to analyze the differences in distribution and trends across genres. By carrying out secondary collocate and n-gram analyses in addition to frequency and primary collocate analyses, the authors collect data providing a better understanding of the use context for neologisms, in a case study of the neologism *K-quarantine*. Finally, they propose a microstructural model for COVID-19 neologisms that integrates the findings of the study, taking the neologisms *Wuhan Pneumonia* and *K-quarantine* as examples.

The results presented in this paper show that comment data prove invaluable for lexicographic description of neologisms: “The value of comment data in lexicographic description ultimately lies in the pragmatic information and the socio-cultural background it provides on headwords and which are not easily seen in existing dictionaries. Moreover, unlike articles, comments are produced by a multitude of commenters and reflect their emotions and stances in relation to the relevant neologisms, providing dictionary users and future generations with fresh, raw examples of real-life language for neologism headwords.” The authors concede, though, that experts need to decide to what extent the politically incorrectness of commenters’ language may be used in dictionaries.

Moving to the lexicographic description of COVID-19 neologisms in Spanish, it becomes evident that the question of a lexeme losing its neologism status by being included in general dictionaries also needs to be discussed. In their paper **Lexicographic detection and representation of Spanish neologisms in the COVID-19 pandemic**, Pedro J. Bueno and Judit Freixa “address the neological process and [. . .] reflect on the various stages of it, from the time a neologism is born until the moment it ceases to be one because it has been dictionarised” (i.e. incorporated into a dictionary). Based on their definition of “pandemic neologisms” and the neological process, the authors give information on their corpus data and data analysis methods before presenting three different groups of COVID-19 neologisms: “non-dictionarisable

neologisms”, “neologisms in the antechamber of dictionarization”, and “dictionarizable neologisms”. They also discuss how some of the neologisms found in their study have recently been added to the *Diccionario de la lengua española* (DLE), the authoritative Spanish language dictionary published by the Royal Spanish Academy with participation of the Association of Academies of the Spanish Language.

The authors point out that the inclusion of neologisms in dictionaries is “dual property acting on a two-fold plane: that of consolidation in use on the one hand, and that of the criteria governing the elaboration of dictionaries on the other”. Some thoughts also go into the different categories that neologisms fall in: short-lived, fleeting ones, and those staying on and becoming fully lexicalized and accordingly recorded in general language dictionaries.

Andreína Adelstein and Victoria de los Ángeles Boschiroli take this discussion a step further in their paper **Spanish neologisms during the COVID-19 pandemic: Changing criteria for their inclusion and representation in dictionaries**, by looking not only into the inclusion of COVID-19 neologisms in (synchronic and historical) general language dictionaries of Spanish, but also into a bilingual English-Spanish dictionary and a Spanish neologism dictionary aiming to cover geolectal variants in six Spanish-speaking countries in Pan-America. The authors describe the different criteria used in the process of inclusion and treatment of the lexemes in those dictionaries starting their study with data obtained by the *Antenas Neológicas Network*, which are “collected exclusively from the written press of the six countries that make up the network”. They concede that “this may be regarded as a limitation in terms of diaphasic variation in relation to pandemic vocabulary, but on the other hand, it guarantees a certain degree of institutionalization, which is an essential aspect when considering the inclusion of new words in a general language dictionary.”

By comparing how different types of dictionaries include/exclude COVID-19 neologisms, they find that there is “a certain degree of overlap of some features which are traditionally thought to be specific to each type of dictionary: [. . .] Dictionaries which, unlike dictionaries of neologisms (which make no claim to finality of stability regarding the place in the language of the items collected), are not restricted to these phenomena or not supposed to collect them, ended up recording ephemeral or witness items, with a very low or null frequency of use.”

In a third perspective on the Spanish language, Magdalena Coll and Mario Barité focus on the inclusion of technical COVID-19 neologisms into a general language dictionary of Spanish in their paper **Specialized voices in the 23rd edition of the *Diccionario de la lengua española*: Analysis of the COVID-19 field and its neologisms**. By analyzing the lexicographic treatment of specialized language neologisms as well as new words beginning with CORONA-, they assess the particularities of the dictionaries in question regarding the incorporation of the new words, as well as the degree of correspondence or complementarity between the last two editions of DLE. The authors demonstrate how “the new additions open up a debate

on the treatment of neologisms in academic lexicography, in a particularly unique scenario”.

Here again, the rapid vocabulary expansion and its subsequent lexicographic treatment throughout the COVID-19 pandemic is seen as an “opportunity for lexicography and terminology researchers”, who should “discuss and propose consistent solutions for the incorporation of scientific and specialized words into DLE and other Spanish dictionaries” and “leave behind vague criteria for incorporating or excluding scientific terms, scientific definitions not easily understood by a regular audience, conceptual inaccuracies, and somewhat erratic assignments of thematic labels”.

In the first paper of the next group in this volume, **How the COVID-19 pandemic is changing the Hungarian language: Building a domain-specific Hungarian/Italian/English dictionary of the COVID-19 pandemic**, Judit Papp looks into ways of compiling a trilingual online dictionary with COVID-19 neologisms using different corpus and dictionary writing tools: “With the creation of the dictionary, my aim is to fill [. . . the] lexicographic gap primarily concerning the Hungarian-Italian language pair and to organize this content in a free online tool (a rich database) that is easy to search and useful for linguists and translators. The third language is English, as the comparison with it is inevitable. [. . .] papers, findings, and results of scientists’ experiments relating to COVID-19 are published in English and this means that English plays an important role in the creation of neologisms. In both Hungarian and Italian, we record a certain number of loans, calques, and adaptations”.

Here, again, the author interprets the high number of COVID-19 neologisms as a sign for the creativity and vitality of a language (namely Hungarian), and discusses how these aspects affect the lexicographic description in an online dictionary (here a trilingual dictionary of equivalents).

Questions of standardization not only arise regarding terminology, but also in connection with general language. In their paper **Coronavirus-related neologisms: A challenge for Croatian standardology and lexicography**, Milica Mihaljević, Lana Hudeček and Kristian Lewis discuss which COVID-19 neologisms collected from media corpora and online sources should become part of general language dictionaries. They distinguish between Croatian neologisms (single and multiword units) and loanwords and loan translations and stress the importance of responding with prescriptive information in their dictionary to the high number of user questions (regarding orthography, morphology, word formation, usage in a sentence and, last but not least, meaning) concerning all types of neologisms.

Their starting point for the lexicographic description of COVID-19 neologisms was the *Glossary of Coronavirus* compiled by a small group of lexicographers with a clearly descriptive intention: “The purpose of the *Glossary* was to meet the needs of Croatian speakers as soon as possible. It usually records terms as they are used and does not give any normative advice. It includes jargon words as well as scientific

terms which entered the general language”. Entries in this glossary were then systematically searched in those corpora that are the basis for the *Croatian Web Dictionary – Mrežnik*, a normative dictionary. The comparison between the differences in the prescriptive vs. the normative approach is informative for other dictionary projects as well.

In Sílvia Barbosa and Susana Duarte Martins’ paper **The neologisms of the COVID-19 pandemic in European Portuguese: From media to dictionary**, we learn about the occurrence of COVID-19 neologisms in the press and social networks and whether and how European Portuguese dictionaries have incorporated them. The authors focus on four candidates: *COVID-19*, *coronavirus*, *pandemia*, and *tele-*, and demonstrate with many examples how these are incorporated into new morphological formations, illustrating how vital the lexical neology process in the domain of COVID-19 in a rather short period of time (2020/2021) actually was.

This study also sheds light on how online dictionaries find different ways of reacting to sudden vocabulary expansion, but also on how the Portuguese language was adapted to the new situation by all its speakers. The authors state (what is true also for many of the examples given in other chapters in this volume): “Only the future will tell whether the creative linguistic phenomenon that emerged from the pandemic will persist in the Portuguese language (namely the loss of the neologism status of particular units while being incorporated in the current language lexicon) or whether it will be a source of occasionalisms circumscribed in time and space while the COVID-19 outbreak lasts.”

Ieda María Alves, Beatriz Curti-Contessoto, and Lucimara Costa present data on Brazilian Portuguese COVID-19 terminology in their paper **COVID-19 terminology and its dissemination to a non-specialised public in Brazil**. Their corpus-based study “aims to detect, analyse and discuss the characteristics of COVID-19 terminology, in particular the role of the adjective *novo* [*new*] in this terminology, the high recurrence of terms in the plural and the resemantisation of some of the terminological units used”.

Their ultimate goal is to create a “terminological dictionary aimed at non-specialised readers in the medical field with little formal education”, in which the terms will be presented onomasiologically. As the intended user group comprises a high percentage of functionally illiterate people, the terms will be defined using plain language. The paper exemplifies the manifold lexicographic problems arising when dealing with new terms from the Coronavirus pandemic in such a setting.

In their paper **Neoterm or neologism? A closer look at the determinologisation process**, Rute Costa, Margarida Ramos, Ana Salgado, Sara Carvalho, Bruno Almeida, and Raquel Silva focus on new lexical units in the Portuguese media discourse and their formation, categorization, and lexicographic description. Especially words formed with *covid-* are collected and analyzed regarding the question “whether these words can be considered neoterms or, on the contrary, if having a term in their formation

corresponds to a false neological intuition. In the latter case, rather than a neoterm, we have a neologism resulting from a process of determinologisation.”

The authors also discuss several issues regarding the inclusion of such “neoterm” in dictionaries, for example their definition and which domain label should be used. In a template proposal for a lexicographic entry, the authors present their ideas and reflect on the “dictionary as a language model,” giving “descriptive guidance” to its users.

In some ways different, but also comparable, problems arise for Sign Language and its lexicographic description, as shown by Mireille Vale and Rachel McKee in their paper **Neologisms in New Zealand Sign Language: A case study of COVID-19 pandemic-related signs**. New signs for suddenly very frequently used new terminology regarding COVID-19 had to be created, conventionalized and disseminated throughout the community of Deaf people in New Zealand. The authors also aim “to explore how and when such neologisms could be entered in the ODNZSL” (*Online Dictionary of New Zealand Sign Language*). The data on signs related to COVID-19 was collected from two sources: signs that were contributed to *NZSL Share*, a web-based platform where users can upload sign videos etc., and signs used by interpreters (e.g. while translating TV briefings on the Corona pandemic). To form the new signs, different strategies were used including “semantic extension; coinage of new words through language-internal mechanisms such as derivation or compounding; and drawing on language-external resources, as calques or direct loans”.

Regarding the lexicographic treatment of such lexical innovations, similar problems to conventional language arise, as signs to be included into a dictionary should be fixed, used over a longer period of time outside the original context and widely throughout the whole Deaf community. Using a crowdsourcing platform like *NZSL Share* seems to be a promising tool to find and spread Sign neologisms that then help to update *ODNZSL*.

In the closing paper of this volume, **Using Wiktionary revision history to uncover lexical innovations related to topical events: Application to Covid-19 neologisms**, Franck Sajous explores how data from current revisions in Wiktionary (here demonstrated with the English and the French versions) can be explored to find candidates for COVID-19 neologisms for inclusion in other dictionaries (in addition to exploring media corpus data), thus enabling lexicographers “to monitor, analyse and report quickly a sudden inflow of lexical changes”. After explaining his methodology (data processing, ranking new and existing headwords, and annotation of headwords), the author presents his results. Here, readers learn about the different contributor types and how existing and new entries are ranked quarterly and annually, as well as on false negatives.

The study is “based on the hypothesis that Wiktionary’s most heavily modified articles can help detect new and existing headwords that are related to topical events”, which could be validated for COVID-19 neologisms, at least regarding the English and French Wiktionary version with very active online communities. It

remains to be seen, however, whether the method described here will be able to detect lexical innovations related to topical events with a smaller impact than the Corona pandemic evidently had.

Overall, the findings of the studies in this volume focus on how lexicographic work regarding COVID-19 neologisms has been done and could be improved, either by exploring corpora and other data more systematically, by incorporating users' expertise into the lexicographic process, or by learning from the lexicographic practice of existing dictionaries. Many authors also stress how strongly lexicographic work was affected by the COVID-19 pandemic and its repercussions on the vocabularies of languages around the world in many different ways, but also how, due to such challenges, steps were taken to improve lexicographic work. We hope that the discussion regarding these and other questions related to lexicography and neology in the context of the COVID-19 pandemic and beyond will continue and that this volume contributes to it in a fruitful way.

We would like to express our gratitude to the editorial board of *Lexicographica. Series Maior*, primarily Stefan Schierholz, who quickly accepted our proposal for this volume into the series. And we thank the authors, at the heart of this publication, for their contributions and trust in us.

Bibliography

Klosa-Kückelhaus, Annette/Kernerman, Ilan (2020): Global Viewpoints on Lexicography and Neologisms: An Introduction. In: *Dictionaries. Journal of the Dictionary Society of North America* 41.1: 1–9.

Klosa-Kückelhaus, Annette/Kernerman, Ilan (2021): Converging Lexicography and Neology. In: *International Journal of Lexicography* 34.3: 277–281.