POSTPRINT

Global Viewpoints on Lexicography and Neologisms: An Introduction

Annette Klosa-Kückelhaus and Ilan Kernerman

Leibniz-Institut für Deutsche Sprache, K Dictionaries

klosa@ids-mannheim.de, ilan@kdictionaries.com

ABSTRACT

This is an introduction to a special issue of *Dictionaries: Journal of the Dictionary Society of North America*. It offers a characterization of neology and describes the Globalex-sponsored workshop at which the papers in the issue originated. It provides an overview of the papers, which treat lexicographical neology and neological lexicography in Danish, Dutch, Estonian, Frisian, Greek, Korean, Spanish, and Swahili and address relevant aspects of lexicography in those languages, presenting state-of-the-art research into neology and ideas about modern lexicographic treatment of neologisms in various dictionary types.

Keywords: neologisms, lexicographical neology, neological lexicography

LEXICOGRAPHICAL NEOLOGY AND NEOLOGICAL LEXICOGRAPHY

Neology constitutes a natural, dynamic, and multilateral part of all living human languages, whether as a reflection of or to facilitate linguistic communication, and lexicographic interest in neologisms dates back to the 1600s, when neologisms started to be included in hard-word

dictionaries (see Barnhart and Barnhart 1990, 1162). There is a vast field of research on 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, as well as combining both), distribution (in general language and in domain-specific language, i.e., 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 understanding of neologisms as applied here includes new words, new multiword units, new elements of word formation, and new meanings of any of them.

In the context of lexicography and neologism, important lexicography-driven or lexicography-oriented aspects need to be addressed, including these:

- how differently, if at all, neologisms should be treated in different dictionary types (e.g., in historical comprehensive ones and those focusing on current usage; in monolingual and 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 well neologisms that are integrated in dictionaries are accepted by the community (issues of rejection of new words and language purism)

- how dictionary users can help find neologisms and convey information about them

- how to deal with grammatical/orthographic/pronunciation variation (descriptive vs. prescriptive approaches)

- how to interoperate lexicographic datasets with online resources and incorporate neologisms into dictionaries (media, formatting, labeling, etc.)

- how to explain meaning with and without encyclopedic information and how to use illustrations and audio-visual media

The papers in this special issue of *Dictionaries: Journal of the Dictionary Society of North America* discuss some of these aspects, presenting

state-of-the-art research into neology and ideas on modern lexicographic treatment of neologisms in various dictionary types.

GLOBALEX WORKSHOP ON LEXICOGRAPHY AND NEOLOGISM

This special issue features eight papers from the Globalex Workshop on Lexicography and Neologism (GWLN),[1] which was held on May 8, 2019 in conjunction with the 22nd Biennial Meeting of the Dictionary Society of North America (DSNA) at Bloomington, Indiana.[2] GWLN included both in-person and online presentations of thirteen papers from academia and industry on issues related to the interaction of lexicography and neology, focusing on a dozen languages: two papers on English and one each on Danish, Dutch, Estonian, Frisian, German, Greek, Hebrew, Japanese, Korean, Spanish, and Swahili. Presenters who were not able to attend the conference delivered their papers and participated in the discussion by means of video conference, broadcast from as far away as Korea, Israel, Greece, and Spain, and as relatively nearby as California.

The workshop was sponsored by Globalex, the Global Alliance for Lexicography, which is operated by the five continental associations for lexicography—Afrilex, Asialex, Australex, DSNA, Euralex—with the goal of promoting international exchange and cooperation on lexicography.[3] Two Globalex workshops had been organized previously as part of the LREC series of international conferences on language resources and technologies[4] in 2016 (Portorož, Slovenia) on the topic of Lexicography and Human Language Technologies[5] and in 2018 (Miyazaki, Japan) on Lexicography and Wordnets.[6] The Globalex workshop in which the papers in this issue originated was the first to be held in the framework of a conventional lexicography conference of one of its member associations.

[1]https://globalex2019.globalex.link/
[2]https://dictionarysociety.com/conference/
[3]https://globalex.link/
[4]http://lrec-conf.org/
[5]https://globalex2016.globalex.link/
[6]https://globalex2018.globalex.link/

The aim of GWLN 2019 was to explore issues related to lexicography and neology for languages across the world, look into their unique characteristics and common denominators, and learn from their similarities and differences with regard to problems, methodologies, processes, and solutions. Being the first such attempt to deal with this domain on a relatively global scale, our approach was cautious and low-profile. The initial plan was to invite around six participants, but the enthusiastic response we received prompted us to increase the number of invitations. This small-scale experiment felt somewhat like scratching the surface, making clear to us that the topic of lexicography and neologisms is highly popular and relevant for numerous languages on multiple lexicographic layers and that researching much further would be worthwhile. Now Globalex hopes to pursue this workshop on an annual basis at conferences of its member associations.[7]

THIS SPECIAL ISSUE

The eight papers included in this special issue concern six European languages, one African language, and one Asian language. Taken together, they reflect the idea that defining what neologisms are, finding them by applying various methods, choosing those to be treated in a specific dictionary, and deciding on the kind of lexicographic information given for them are global challenges to lexicography.

In the first paper, **Lars Trap-Jensen** discusses the inclusion of language-internal neologisms and Anglicisms in a monolingual general dictionary of contemporary Danish ("Language-internal neologisms and Anglicisms: Dealing with new words and expressions in *The Danish Dictionary*"). The data presented about the extent of pressure from English and its impact on the vocabulary (as covered in the dictionary) exemplify similar situations for languages other than Danish. The author shows that a little over half of the neologisms recorded in *The Danish Dictionary* are language-internal compounds and derivations based on existing Danish words and morphemes, while direct borrowings from other languages (predominantly English) make up more

---

[7]The second GWLN is scheduled to take place in conjunction with the Euralex 19th International Conference at Alexandroupolis, Greece in September 2020 (see https://globalex2020.globalex.link/gw-euralex2020/).

than a quarter of all neologisms. Blends (borrowings in combination with Danish words or morphemes) are the third largest group, and only three percent of all neologisms in this dictionary are loan translations. Trap-Jensen's conclusion that "[n]eologisms are important to keep a language dynamic and adaptive to an ever-changing world" and that "[a]t present, English is the dominant lingua franca of the world, and it comes as no surprise that this is reflected in the vocabulary when words for new things and phenomena come into the language" is relevant for all other languages as well.

In Dutch, the language covered in **Vivien Waszink**'s "Neologisms in an online portal: The Dutch *Neologismenwoordenboek* (NW)," we find similar examples for newly coined compounds or derivations, new meanings, new multiword units, and loanwords from English. Because NW is not a general dictionary but one covering neologisms exclusively, however, the criteria for inclusion of new words and new meanings are different, as are parts of the lexicographic information given (e.g., a stress on etymological information which not only gives the word's root but tells a little story about "who has coined a word or who has made a neologism popular" if possible). Also, more ephemeral neologisms are included (in comparison with a general dictionary of contemporary Dutch), regardless of their possibly short life cycle. In order to find candidates for inclusion, this project combines corpus-linguistic methods with editorial data evaluation in the specifically designed system called Neoloog.

Many neologisms in Frisian, a lesser-used language spoken in the Netherlands, are borrowings from Dutch, as **Hindrik Sijens** and **Hans Van de Velde** show in "The formation of neologisms in a lesser-used language: The case of Frisian." The authors discuss which neologisms should be included in a bilingual Dutch-Frisian online dictionary, dealing with one major issue: there are far fewer neologisms formed in Frisian than in Dutch, the dominant language of the Netherlands, so "for many Dutch and international concepts, there are no Frisian equivalents." Often, speakers codeswitch to a Dutch word (as Dutch and Danish speakers tend to do with regard to English, as reported in the papers by Trap-Jensen and Waszink, this issue), so that it is not always clear whether the Dutch word is indeed a borrowing that should be covered in the dictionary or a nonce word. Overall, more than ninety percent of

neologisms in Frisian are borrowings from Dutch, and many of those have been "naturalized," that is, "adapted according to Frisian phonetic, phonological, morphological or orthographic system." The findings are based on Frisian corpora and Google searches, but plans for the development of a neologism platform as part of the lexicographic workflow are provided.

Estonian is also one of the smaller languages in Europe, but one with findings about neologisms that differ from those about Frisian, as **Margit Langemets**, **Jelena Kallas**, **Kaisa Norak**, and **Indrek Hein** show in "New Estonian words and senses: Detection and description." The authors present data on semi-automatic new word detection and lexicographic treatment of new words and meanings in an Estonian language portal. Many of the neologism candidates found are compounds and derivations, that is, language-internal neologisms, as "Estonian is a predominantly agglutinative, highly inflected language with a productive and flexible morphological derivation system." Because they are semantically transparent, not all of these can or need to be included in a dictionary, so further development of the tools used for automatic language processing is needed, as well as means to detect new meanings and multiword units. The authors also note that it is "important to automate the tracking of particular neologisms in order to provide users with up-to-date information on their meanings and usage," which of course is especially relevant for online publication of dictionaries.

While the dictionaries introduced in the first four papers in this issue cover general language, **Gilles-Maurice de Schryver** takes a close look at neologisms in terminology in "Linguistics terminology and neologisms in Swahili: Rules vs. practice." For this paper, linguistics terminology was extracted semi-automatically from Swahili language and linguistics textbooks, then analyzed with regard to either source languages (e.g., English *sentence* > Swahili *sentensi*) or word formation processes (e.g., *matamshi* 'pronunciations, speech' from *kutamka* 'to pronounce, to articulate'), and finally, examined as to how these neologisms are treated in existing reference works. As "[l]anguage and especially modern linguistics being a relatively recent field of knowledge for which terminology has had to be developed in Swahili, … quite a number of Swahili language/linguistics terms are still novel enough to be called neologisms." Only a little more than half of neologisms in linguistics

terminology are covered in a general language dictionary of Swahili and only sixty-one percent are included in a Swahili-English linguistics terminology list. In conclusion, language-internal neologisms should be treated in dictionaries besides the borrowings, as they are often "the result of semantic specialization."

Tracking neologisms and including them in an online language portal has a long tradition in South Korea, where almost twenty-five years ago the investigation of Korean neologisms was initiated by the government, as **Kilim Nam**, **Soojin Lee**, and **Hae-Yun Jung** explain in "The Korean Neologism Investigation Project: Current status and key issues." Although dictionary users contribute to keeping the information in a Korean language portal up to date to some extent, the authors stress "the crucial role of expert lexicographers" for finding and describing neologisms. By comparing findings from the Korean Neologism Investigation Project with data from four monolingual general language dictionaries of Korean, the authors demonstrate that "prescriptive dictionaries … tend not to include neologisms … [while] user-generated content dictionaries tend to include high-frequency neologisms." In order to facilitate lexicographic decisions on the inclusion or exclusion of neologisms, "usage trend investigations" to determine which neologisms survived are proposed.

In **Judit Freixa** and **Sergi Torner**'s paper "Beyond frequency: On the dictionarization of new words in Spanish," other possible criteria for the inclusion of new words in dictionaries are discussed (e.g., morphological features, semantic transparency, corpus frequency). The authors introduce a combination of "change in frequency of neologisms through time with speakers' perceptions about their novelty" as a criterion for their "dictionarization," that is, their incorporation into a dictionary. Speakers' perception of the novelty of words is "shaped by different parameters …: the impact of novelty, frequency of use, and non-regular formation of the unit, … type of neologism, use in special domains, and register of use." The findings show that neologisms formed by irregular mechanisms or loans for which Spanish words already exist, for example, do not need to be included in general dictionaries for native speakers. Only in combination with observing frequency development can good decisions on the dictionarization of neologisms be made by lexicographers.

In the final paper, the focus lies once again on a specific type of neologism, this time of genericized trademarks. **Anna Vacalopoulou** explores "Criteria for the inclusion of trademarks in general language dictionaries of Modern Greek" and compares the treatment of trademarks in general language dictionaries of English and Modern Greek, relying on corpus evidence. Lexicalization criteria for the inclusion of trademarks are their use with lower case spelling in corpora, morphological adaptions, demonstration of productive capacity, attestation of plural forms, and metaphorical use. "The exploration revealed a distinct increase in inclusion of this type of neologism" in dictionaries of Modern Greek. This reflects a general trend of a growing number of such neologisms in modern languages (e.g., in Dutch, where quite a number of neologisms from trademarks are attested, as Waszink points out, this issue).

For various reasons, not all of the languages and topics treated at the workshop are represented in this special issue of *Dictionaries*. Consequently, English, German, Hebrew, and Japanese neologisms are not treated directly here. At the workshop, however, issues on English neologisms were presented in Katherine Connor Martin's "A system for evaluating multiple data inputs to prioritize neologisms for inclusion in dictionaries" and Erin McKean's "Using the Hypothes.is web annotation tool for neologism collection," both of which focused on ways of detecting neologisms in the context of their particular projects (Oxford dictionaries and Wordnik, respectively). Kathrin Kunkel-Razum discussed how German neologisms are detected for inclusion in print and online dictionaries of Duden publishing house in "New words for the *Duden*." For Hebrew, Noga Porath presented examples of neologisms recently found by editorial detection, by crowd sourcing, or by the Academy of the Hebrew Language for an online dictionary in "Adding neologisms to the Hebrew online dictionary *Rav-Milim*." Teruaki Oka, in "New words in Japanese and the design of *UniDic* electronic dictionary," described the specific challenges of neologism detection in Japanese corpora and how the findings are represented in a dictionary database.

Overall, diverse answers to the questions raised above are proposed in this special issue, most importantly regarding how to integrate different types of neologisms into different types of dictionaries. Questions about language purism and consequent rejection of new words are also addressed in some papers and for some dictionaries. Finally, ideas on how dictionary users can help with finding, evaluating, and informing

about neologisms and on how to deal with neologisms that are no longer new or no longer used are presented. Other questions remain unanswered, such as how to explain the meaning of neologisms with and without encyclopedic information and how to use illustrations and audio-visual media. The discussion will continue—we hope this special issue of *Dictionaries* will contribute to it in a fruitful way.

REFERENCE

Barnhart, Robert, and Clarence Barnhart. 1990. The dictionary of neologisms. *Wörterbücher - Dictionaries - Dictionnaires. Ein internationales Handbuch zur Lexikographie*, Vol. 2, edited by Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand, and Ladislav Zgusta, 1159–66. Berlin and New York: de Gruyter.