

Gilles-Maurice de Schryver (ed.) *A way with words: recent advances in lexical theory and analysis—a Festschrift for Patrick Hanks*

Menha Publishers, Uganda, 2010, 384 pages. ISBN 978-9970-101-01-6

Shin'ichiro Ishikawa

© Springer Berlin Heidelberg 2014

1 Outline

This book is a festschrift for Professor Patrick Hanks presented on his seventieth birthday. Hanks is widely known as an editor of the Hamlyn *Encyclopedic World Dictionary* (1971), the *Collins Dictionary of the English Language* (1979), the *Collins COBUILD English Language Dictionary* (1987), *The New Oxford Dictionary of English* (1998), *A Dictionary of Surnames* (1988), *A Dictionary of First Names* (1990), and *Dictionary of American Family Names* (2003) as well as the author of more than one hundred academic books and papers on lexicography.

Reflecting Hanks' widespread interests, this book includes 20 excellent papers discussing various facets of modern dictionary making. The book begins with an informative introduction by Gilles-Maurice de Schryver. The editor of the book outlines Hanks' life and academic contributions and explains his theory of norms and exploitations, which concerns two kinds of rules governing the normal and creative use of phraseology, as well as his corpus pattern analysis (CPA), which tries to define the behaviour of a verb from its patterns, namely, its valency structure and its collocational preferences.

The main body of the book comprises three parts, which focus on three key aspects of lexicography: theory, computation, and the practice of dictionary writing.

S. Ishikawa (✉)
Kobe University, Kobe, Japan
e-mail: iskwshein@gmail.com

2 Theory of lexicography

Part 1, which includes five papers on theoretical aspects and the background of lexicography, begins with the last and unfinished paper by the late John Sinclair. Emphasising that a single word is no longer regarded as the principle unit of meaning in language use, Sinclair analyses the behaviour of the verb *sever* by paying attention to its collocating prepositions, colligations, and semantic preferences so as to reconsider what the dictionary definiendum can be.

Yorick Wilks pays attention to the fact that pre-determined semantic preferences and restrictions can be broken in language use and insists that this kind of word sense extension is a part of our natural language understanding system and it should be treated as something general. His aim is to clarify the complex role of knowledge in language understanding.

James Pustejovsky and Anna Rumshisky also discuss semantic extension with a special focus on verbs. They insist that the degree of meaning extension, in other words, that of metaphoricity, is determined by three kinds of formal processes operating on the predicate: (a) generalising the type of the argument of the predicate, (b) changing the argument structure and relative prominence of arguments, and (c) abstracting the core meaning of the verb.

Igor Mel'čuk illustrates how lexical government should be described in his framework of explanatory combinational dictionary (ECD), which is based on his meaning-text linguistic model. In ECD, the government of a lexeme is analysed from its idiosyncratic pattern concerning all of the deep syntactic, surface syntactic, and morphological levels.

David Wiggins, in his somewhat philosophical essay, explores the paradox of analysis and synonymy, which clearly concerns the act of definition in lexicography. If a definition (what analyses) is completely synonymous and literally identical with the headword (what is analysed), the act of definition or that of analysis is practically nothing; while if they are not in accordance, it means that the act is wrong.

3 Computation and lexicography

Part 2 of the book includes seven papers exploring the computation of lexical relations. Modern dictionary making has moved far away from the traditional pen-and-paper-based editing style, as the extensive use of computers and machine-readable corpora for language analysis is now central to the practice.

Kenneth W. Church discusses how corpora should be compiled, in other words, the problem of balance or size. It is generally believed that a small but balanced corpus is better than a large but unbalanced corpus, but based on his own experience, Church notes that an unbalanced, repetitive, and formulaic newswire corpus was more useful for editing a thesaurus and spelling corrections. Church also examines the value of internet search engines.

When determining the number of headwords included in dictionaries, one needs to identify the total number of concepts embodied by existing words and phrases.

Gregory Grefenstette presents an interesting estimation for this purpose. He focuses on single nouns and two word NP combinations and, based on a web corpus search, estimates the number. He also mentions the unreliability of web data and Google counts.

Identification and disambiguation of the meaning of a lexeme are one of the most challenging tasks for lexicographers. David Guthrie and Louise Guthrie attempt to predict the semantic category of the nouns (e.g. abstract, concrete, animate) by their collocating adjectives using two mega corpora. They report that the accuracy rate of their prediction was 70 %.

Alexander Geyken examines whether the light verb + noun combinations presented in German dictionaries are statistically salient in two mega corpora. Interestingly, salience was observed in the 1 billion token corpus, but not in the 100 million token corpus.

Sketch Engine, which produces a collocation pattern of a lexeme in a very readable way, has been used more widely in dictionary making. Karel Pala and Pavel Rychlý examine the lexical profile of Czech words obtained from the Sketch Engine and propose several ideas to correct identified errors.

Although one of the Hanks' outstanding contributions to lexicography is CPA, what matters is its reliability. Silvie Cinková, Martin Holub, and Lenka Smejkalová conducted an experiment probing the degree of inter-annotator agreement in pattern-tagging and report that the overall result was good, though they find several problems caused by confused judgment between a norm use and an exploited use or among marked, unmarked, and undecidable uses.

Elisabetta Jezek and Francesca Frontini reexamine the relation between the verb arguments' semantic types and lexical sets filling different argument positions in Italian and insist that the mismatch between them can be well analysed with Hanks' CPA. They also mention the need for a kind of pattern bank as an extension of a pattern dictionary.

4 Practice of dictionary making

Part 3 of the book includes seven papers discussing aspects of the lexical analysis and dictionary writing. The final part of the book begins with a paper by Rosamund Moon. She focuses on the phraseology of *spring to mind*. Although its presentation varies according to different dictionaries, she examines its frequency, forms, and collocates based on the Bank of English. She reveals that *spring to mind* is a dominant realisation of many related expressions such as *spring [come, leap jump, rush] to mind* and *bring [call, recall] X to mind* and identifies its discourse function.

Sue Atkins compares the DANTE (Database of Analysed Texts of English) as a new database for lexicography, which is based on a 1.7 billion words corpus linguistically analysed with Sketch Engine and covers 42,000 lexical entries, with the famous FrameNet, which annotates 150,000 corpus sentences and defines 1,000 frames and 10,000 lexical entries. She emphasises that the information available only in DANTE, such as secondary grammar (subclasses of nouns, verbs, and adjectives), multi-word expressions including collocations, phrases, chunks,

collocates, itemisers (e.g. *buckets of X* or *a slice of X*), and usage markers, is important for linguistic analysis.

As mentioned, modern lexicography has been increasingly depended on computers and oriented for the automatic processing of language data. Adam Kilgarriff and Pavel Rychlý explore the possibility of using the Sketch Engine as an infrastructure and conducting a semi-automatic dictionary drafting (SADD), which they say to allow the computer to do much of the footwork of varied corpus pattern analyses and to assign a new corpus instance semi-automatically to particular senses. As a word sense is eventually a group of specific lexical instances in use, computer-based examination of a lexeme's collocations, grammar patterns, text domains, and semantic category enables us to group the corpus lines according to similarities in forms and meanings and identify the behaviour of the lexeme.

Is dictionary making a theory or just a practice? This topic is touched upon by Paul Bogaards, who analyses the corpus of *International Journal of Lexicography* compiled by the editor of the book and surveys what leading scholars think of the topic. According to Bogaards, many assume that there already exists some theory in lexicography, though its content is largely unclear, while others explicitly negate the existence of a theory of lexicography. The author says, as there is no single medical theory, there can be no single lexicography theory; and if it exists, it should cover various aspects of dictionary making such as the user, data, and the user's access to the data.

Then, how is a dictionary made in reality? Mirosław Bańko's unique paper describes in great detail how he compiled a COBUILD-like Polish dictionary literally from scratch. He summarises the merits of COBUILD and explains how he has adjusted its editing policy for a dictionary of a language quite different from English. He also mentions that his dictionary has influenced many following dictionaries published in Poland.

When discussing dictionaries, we should not forget varied specialised dictionaries. Jonathon Green discusses the collection and analysis of argots or French slang.

The last paper in the book is that by Michael Rundell, who pays attention to the concept of elegance in dictionary entry writing. Although many dictionary critics say that some entries are linguistically and/or lexicographically elegant and others not, what they mean is not necessarily clear. Rundell discusses the issues such as how to avoid redundancy, how to compress content, whether examples, most typically COBUILD-like examples, are really needed or not, whether lexicographers need to increase or decrease the number of word senses, how to present those senses (based on the history, frequency, or semantic coreness), how to deal with multi-word units, whether full-sentence definitions are good or not, and whether lexicographers need to pursue consistency in word definitions.

5 Conclusion

As briefly summarised above, this book includes papers on a range of topics related to lexicography. Although this is clearly an attraction of the book, at the same time, there are some inconsistencies across the different papers, some of which might

confuse readers. For example, some papers have abstracts, whereas others do not. Some papers are structured in sections, but others are not. Some papers are casual essays discussing general issues in lexicography, while others deal with highly specialised topics. It should be noted that a few papers dealing with theoretical analysis of syntactic structure may be difficult to understand without sufficient prior knowledge on the readers' side. The book has neither an index nor a glossary, which might help readers better understand the content of the various papers.

However, these points do not depress the outstanding value and merit of the book, which is an “essential reading for everyone interested in meaning, the lexicon, dictionaries, and corpus linguistics”, as the editor states. This book also serves as an attractive alternative to traditional textbooks on lexicography (e.g. Landau 2001; Atkins and Rundell 2008; Svensén 2009).

References

- Atkins, B.T.S., and M. Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Landau, S.I. 2001. *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.
- Svensén, B. 2009. *A Handbook of Lexicography: The Theory and Practice of Dictionary-Making*. Cambridge: Cambridge University Press.