

3.1 Researching Lexicographical Practice

Lars Trap-Jensen

Chapter Overview

Dictionary Conceptualization	36
Designing the Database	37
Describing the Linguistic Data	38
Dictionary Writing Systems	43
Data Access and Presentation	44
Finding the Dictionary – the Future	46

The computerization of work routines that the world has witnessed over the last couple of decades has changed the lives of many people, but the effect it has had on dictionaries and on the lexicographer's daily life is all-embracing and difficult to overstate. In this chapter, we look at the various stages involved in dictionary-making and some of the decisions that the lexicographer is faced with in the process.

It should be noted that even if some of the issues are general and shared by different types of dictionaries, others pertain to just one particular type. Monolingual dictionaries are obviously different products compared to bilingual ones, and making a dictionary is different from making an encyclopedia, terminology or even a telephone directory, even if they all must be considered lexicographical products. In the following, the focus of attention is, unless stated otherwise, on monolingual dictionaries for native speakers.

1 Dictionary Conceptualization

While many things have changed dramatically in lexicography, the planning phase is arguably one of the areas that has been least affected, and yet it is perhaps the most important one. It is during this phase that crucial decisions about the database structure and its inventory must be made, based on an analysis of the intended users and their needs. These decisions condition how the data in a later phase can be presented to the end-user and how it can be re-used in other applications.

One thing that has changed is that lexicographers today are less inclined to have one specific product in mind when they build their dictionary database. Over the last decades publishers have spent much effort in unifying their dictionary resources and standardizing the information contained in each element in the database. Instead of offering a range of independent dictionaries, each with their own specific list of entry words, inflectional information, synonyms, style labels, etc., most publishing houses now have one central database from which individual dictionaries can be produced by extracting the desired combinations of information types needed for a particular lexicographical product. From the publisher's point of view, this solution gives them two important advantages. First, it makes maintenance easier, as updates that are made in one element in the database immediately feed through to all the dictionaries in which that element is used. Second, it enables them to refine the range of dictionaries offered, as they can extract different combinations of information types to suit the needs of a specific user group. For the user, it means that they are more likely to recognize a distinct flavour of a particular publishing house's products, and perhaps a sense of familiarity if they buy more than one product.

From a lexicographical point of view, what has happened is that the production of lexicographical data has become more clearly separated from the presentation of the data to the user. Today, many dictionaries are available both as classical paper products (although sales are rapidly declining) and in electronic form. Digitally, they may appear as CD-ROMs, as online versions and as apps for smart phones and tablet computers – and even integrated with other products and applications. The latter includes lexicographical data that is utilized as a resource but is mostly either invisible to the user, as the data used by spell-checkers in word processing programs, or only becomes visible when activated, such as the dictionary definitions found in e-readers that show as pop-ups when users click on a word. More will be said on this in the last section of the chapter.

2 Designing the Database

Lexicography involves a lot of decision-making: How many words should be in the dictionary and by what criteria? What types of information are relevant for the intended target group? Does the intended target group coincide with the actual user group, and if not, does it matter? What is the best way to explain a particular word meaning to the reader? The answers to these and a good deal more are not necessarily easy to provide beforehand, but they are important for the way the database should be built. A database designed to meet future requirements for other dictionaries or publication channels should anticipate as many aspects as possible in the early stage of the process.

To take an example: in a dictionary that is going to appear as a concise paper dictionary, it may be appropriate to use abbreviations, whereas the online version will have the full forms. However, not all abbreviations have a one-to-one expansion: *adj.* refers sometimes to 'adjective' and sometimes to 'adjectives', *bot.* can unfold as 'botany' or as 'botanical'. It is likely that a simple list of abbreviations and their expansions will not do. Instead, all the different possibilities must be taken into account and a special field or an attribute should be available in the database to show how a given word form is presented as a full and abbreviated form respectively.

Another example is morphological information. In a bilingual L1-L2 dictionary, morphological information about headwords is not necessary, as it can be assumed that the users know how the words are inflected in their native language. If the same list of headwords is, at some later point in time, used in a different dictionary, such as a learners' dictionary, no such assumption can be made and the morphological information will have to be produced if it is not in the database from the outset.

For definitions, it is not recommended to use the same wording in a technical dictionary as in an encyclopedia, not to speak of children's dictionaries. For that reason, the database may well include several versions of the same definition to be used for different user groups. Even within the same dictionary, two versions could be offered: a short definition for quick reference and a more elaborate one for users who prefer an encyclopedic explanation with attention to detail.

More could be added to the list of examples: information about pronunciation in either phonetic notation or as sound clips, images and video clips, syntactic and encyclopedic information, quotations and other language examples are all information types that are important to store in the database. They may not all be relevant for publication in one and the same dictionary, but it is advisable to store the information in a central base from where it can be easily retrieved.

In some cases, it may even be practical to include elements in the database that are not ever going to be shown to the end-user but which can be useful for other purposes. In a dictionary project that began in the early 1990s, The Danish

Dictionary, it was decided to include information about the nearest superordinate word (the *genus proximum*) and about subject domain if at all possible. This information was systematically entered by the editors throughout the compiling period but was not used at all in the printed dictionary. It was, however, very useful when, later, the dictionary data was used to build a Danish wordnet (on the model of Princeton WordNet) and to compile a Danish thesaurus (Pedersen et al. 2009, Lorentzen and Trap-Jensen 2011).

Technically, there is a wide range of software solutions available. Some lexicographers and publishers prefer relational databases, others XML bases, and both types exist as proprietary commercial products and as open source products. No more will be said about soft- and hardware, but it should be stressed that the notion 'a central database' is used here as a broad cover term. An actual implementation often involves several databases. The main point is that the overall architecture should be such that the bases are designed to function as a conceptual unit, linked to each other via unique ID numbers.

Apart from defining what elements to use in the dictionary, it is important in the planning phase to prepare a manual or style guide that tells the lexicographers about the inventory of elements and how they should be used. A style guide is especially vital for larger projects with a staff of considerable size, and for long-term projects that have to account for some degree of staff turnover. It is an obvious boon for training new editors and helps to secure a uniform final appearance. Style guides are project-internal tools and as such they vary greatly from project to project, ranging from rough principles (*be brief and to the point; don't use brackets and exceptions if you can avoid them*) to very specific instructions for certain elements (*use a maximum of four synonyms; only describe syntactic patterns with ten corpus examples or more; in metatext, use only words from the defining vocabulary*). A style guide that carefully records all the principles and conventions defined in the planning phase, supplied with the revisions and adjustments made during the compiling process will ultimately capture what in the end gives the dictionary its own characteristic style and personality.

3 Describing the Linguistic Data

After the initial planning phase, where all the general decisions are made, it is time to consider the object of description, the linguistic data. This is an area that has undergone a dramatic development over the last decades, both in the methods used and in the resources available. The achievements within the field of corpus linguistics have produced a range of tools that lexicographers use to establish a sound empirical basis for their linguistic description. Corpus linguistic methods are employed at almost every stage of the dictionary entry: lemma selection, lexical variants, inflection, collocations, valency patterns, set

phrases, compounding and derivation. This interesting topic is explored in further detail in Chapter 4.1. Here, we will start by taking a closer look at the empiricist position and ask whether it is as justified as many lexicographers are inclined to think.

3.1 Prescription or Description?

Historically, the view that dictionaries should reflect the language of all its speakers cannot be taken for granted. In the nineteenth century and earlier it was widely held that, because of the important educational role of dictionaries, they should be normative in the true sense of the word: serving as an exemplary model for their users. Consequently, headwords and examples were excerpted from texts written by respected, canonical authors of their time. A well-known case in point is the *Dictionnaire de l'Académie française*, which set an example for a number of national dictionaries in the eighteenth and nineteenth centuries. To illustrate, one of the pioneers behind the Dictionary of the Royal Academy in Denmark (Langebek 1740) claimed that there was no room in the dictionary for:

All coarse, rude and lecherous words and phrases which contradict decency . . . for they need not be known to those who do not appreciate it, and those who do will surely get to know them anyhow.

And a hundred years later, the editor of the most popular Danish dictionary of the time wrote in his preface (Molbech 1859: viii):

Even the most frequent use of a newly formed word, especially in colloquial language, renders it no authority or proof of usability in pure speech and good style, nor of its admittance into a dictionary as long as it offends the cultivated ear and the delicate language instinct.

There are notable exceptions to the normative tradition, but even so it is not until well into the twentieth century that it became generally accepted for dictionaries to reflect the language community taken as a whole. No doubt, the greater availability of texts beyond the professional works of authors and journalists played a role in paving the way for the descriptive view dominant in the latter half of the twentieth century.

Most lexicographers today accept the descriptive role of dictionaries and prefer to see their own role as objective observers of linguistic facts, but there are areas of lexicographical practice that fall outside the scope of description. Whenever the normative role of language is involved, an element of authority

and language policy is present. The orthographic forms of the headwords in a dictionary are in many countries regulated not directly by the practice of the language users, but at most indirectly via an official body that has been given the formal authority to decide how words are to be spelled. Other countries, such as the United Kingdom, have no such body but a *de facto* norm is set by one or two dictionaries which are widely recognized and followed by the educational system and by central authorities.

3.2 Lemma Selection

Another area where normative aspects are involved is lemma selection, clearly illustrated by the above quotations. A strictly descriptive approach would involve ranking all the words of a well-balanced corpus and mechanically selecting the most frequent ones until a given cut-off point, determined externally by the size and resources of the dictionary project. In itself, it is no trivial matter to decide what constitutes a word or, more precisely, a lexical unit, but we cannot go into the details here. However, very few dictionaries build their headword list in this mechanical way, as the frequency principle would inevitably produce a number of undesired headwords. Most obvious examples are proper names, which occur frequently in corpus texts but are for the most part uninteresting for a general language dictionary. Admittedly, there are exceptions, such as proper names with a metonymic function (*The White House, Mecca*), names that are part of multiword expressions (*Adam's apple, Rome wasn't built in a day*) or culture-specific names that require explanation (*American idol, the London Eye*).

Apart from proper names, compounds and combining forms (*long-tailed, long-haired, long-eared*) are examples of words that are often frequent in texts but are not always obvious lemma candidates. They are often semantically transparent and thus predictable from their components. It should be noted, though, that the process of compounding and derivation is language specific but in languages where the process is productive (which is in general the case for Germanic languages, although less pronounced for English) the result may be a large number of often trivial compounds. In many instances, therefore, the user is better off being able to look up less frequent simplex words which cannot be decoded immediately.

Conversely, the descriptivist model would most likely lead to accidental lexical gaps in dictionary coverage. Parts of a language's vocabulary are made up of closed sets of lexical items, and most people would find it odd if they could only look up some of the months of the year or all the days of the week except *Tuesday*. For systematic reasons, the solution would be to include all the members of the set no matter if, by chance, one or two were not sufficiently represented in the corpus to warrant their inclusion. Even if absence from the

corpus is non-accidental, inclusion may be worthwhile after all. A case in point is the chemical elements, some of which are undoubtedly better known and used than others.

Another problem with lemma selection is the difficulty involved in defining what lexical units belong to a particular language. We have seen that the descriptive approach attempts to reflect the language of the whole language community. But how exactly is a language community delineated? There is no doubt a common core of words that are known to all speakers of English. As one moves away from the common core, however, the vocabulary of individual speakers becomes gradually less concordant. Due to differences such as age, education and housing history, the linguistic experience of a middle-aged engineer from Manchester is different from that of a university student in Cardiff, which is again quite different from a fisherman from Aberdeen. The engineer knows many technical terms from his field of speciality, the fisherman is familiar with the words associated with fishing gear and navigation at sea, and the student probably knows many slang words and informal expressions the others don't, apart from the special vocabulary associated with her subject of study. Due to differences in personal life and linguistic experience it is unlikely that any two speakers of a language have exactly the same stock of words at their disposal. How should the dictionary deal with this? Should it include all the technical terms from subject fields, and all slang, jargon and informal expressions? Ideally, perhaps yes, especially in an electronic dictionary where physical space is irrelevant. In practice, lexicographers are forced to decide on priorities, in which case it is important to realize who is the intended target group of the dictionary. For a learner's dictionary, the users can be expected to look up slang and informal expressions more often than special terms from the fishing trade, and they are also more likely to come across the special words used in linguistics and language pedagogy than words belonging to engineering.

When it comes to regional language, the practice of most dictionaries is to leave out genuine dialect words that are rare outside the geographical area where the dialect is spoken. Instead, these are included in special dictionaries devoted to that particular dialect. Somewhat more controversial are words from other languages that appear in the corpus texts. In the English-speaking world this is perhaps not as controversial an issue as it can be in other countries and languages around the world where the dominant influence of English as a global language is felt. Because of the status associated with the language, English words and expressions appear quite frequently in otherwise 'pure' (Spanish, Czech, Swedish, etc.) contexts. The lexicographer must determine whether to treat these items as loanwords that need explanation like any other word turning up in a corpus with a sufficient frequency, or if they should be interpreted as instances of code switching they can safely neglect. There is no simple answer to this problem, and the lexicographer must in each case carefully

analyse if the item shows signs of integration into the surrounding language, for example in the way the word is pronounced, inflected or used syntactically. The more established the word is in the new language, the more reasonable it is to include it in the dictionary. One should, however, be aware that practice varies significantly, as every country and language have their own cultural and political contexts and traditions. This can be a highly sensitive matter, especially in areas where a minority language has been historically dominated by a larger, perhaps colonial language.

3.3 Language Policy

Dictionaries and language policy can play important and active roles in contributing to the cultural identity and self-understanding of a young nation. Think of the status of Russian in the Baltic states, or of the role of dictionaries for minority languages such as Frisian, Basque, Irish, Sami, etc., where the attitude towards loanwords from the dominant language easily comes to carry political overtones. What in one context is viewed as linguistic puritanism may in another be interpreted positively as a sign of pride in the local language. In some countries much effort is spent in coining new words and expressions in the local language in order to avoid the influence from English or another dominant language. Such an undertaking is, of course, politically rather than linguistically motivated. From the language's point of view, it doesn't matter if the Icelandic word for a female flight attendant is *stewardess* with an English loanword or *flugfreyja* (literally 'flight-Freya', after the goddess of love in Nordic mythology) or if the English word *computer* is used instead of the Icelandic coinage *tölva* (a contraction of *tala* 'number' and *Völva*, a soothsayer mentioned in the younger Edda). What is important, however, is that the language policy is actively supported by the population, whatever direction it takes. Otherwise it may lead to the absurd situation where the dictionary lists one set of words but you hear a totally different set when you visit the local pub.

Even if a solution is found for the descriptive problems discussed here, and even if the achievements of corpus linguistics have indisputably made life easier for the lexicographer in many ways, it should not be forgotten that a substantial amount of data found in dictionaries still cannot be verified empirically. Whether *scent* and *perfume* are synonyms and what their most appropriate equivalents are in French, or whether *scumbag* should be labelled 'informal', 'derogatory' or 'slang' are not questions that can be answered by checking against empirical evidence in a corpus. They are the result of the lexicographer's evaluation based on his or her professional skill and linguistic perception. Writing a precise, informative and elegant definition is still an area where man is superior to the computer.

Finally, the role of and limits to the use of corpora have been questioned in recent years. For a long time, corpus frequency has been unrivalled as the dominant criterion for lemma selection. But one could also ask: can it be taken for granted that the most frequent words are also the words that users want to look up? Traditionally it is a question that was difficult if not impossible to examine empirically. With the arrival of e-dictionaries, log-file analysis can provide valuable data. Those studies that have been carried out (Bergenholtz and Johnsen 2005; de Schryver et al. 2006) suggest that there is in fact little correlation between corpus frequency and look-up frequency. On the other hand, there is still a long way to go before we can predict which words will be looked up in a dictionary and which ones will not. It is simply an area where we have too little knowledge at present. Undoubtedly, it is a field that will attract more attention, not least because corpus-driven dictionaries are being put under pressure from user-driven tendencies. Future dictionaries may well use no-match lists from the log-files rather than corpus frequency as the main criterion for lemma selection.

4 Dictionary Writing Systems

Another area where computers have made life easier for lexicographers is the software they use for entering the lexicographic data into the database. Dedicated dictionary writing systems (DWS) help build the data structure and secure data consistency. They are designed to implement some of the decisions that would formerly be part of the style guide. By creating a Document Type Definition (DTD) or, more recently, an XML schema for the document in which the dictionary is being edited, the lexicographers can specify everything related to the document structure: what elements can be used, in what order are they allowed to occur, which elements may be used recursively, what content is possible (characters, images, sound or video clips), and what attributes an element can have. If an editor makes a mistake in attempting to store the article document, he or she is notified immediately and presented with the possible causes of schema violation.

Cross-references are another traditional source of errors in dictionaries that can be handled by a DWS that binds and automatically tracks the source and targets of a reference. Again, if an editor deletes or changes either of the two, he or she will be notified and can take appropriate action.

Most DWSs offer various other features, such as: advanced search and statistics, preview settings, export or publishing modules, integration or interoperability with other bases and multi-user set-up. If the DWS has a login function, it can be used as a tool for the project management to keep track of article production and workflow in the various editorial phases. DWSs may be developed

and tailored to meet the exact needs of a specific project, but there are also several off-the-shelf products available on the market that are sufficiently flexible to meet most customization needs.

5 Data Access and Presentation

As mentioned earlier, there is a growing demand for dictionaries to be available in various channels and on several platforms. This implies that their contents must be presented to the user in different ways, as the possibilities in a printed dictionary are very different from those of an internet browser – the use of hyperlinks and audio/video clips are obvious examples. Likewise, the limited size of the screen constrains what can be displayed on smartphones and other mobile devices in comparison with a 24-inch desktop monitor. If the structure of the dictionary has been devised with sufficient care, it is possible to take the differences into account in the publishing phase.

The function and aesthetics of layout and typography in general belong to a long and well-established tradition with obvious consequences for lexicography. However, the readers are encouraged to explore for themselves the wealth of literature on the subject as no more will be said about it here. Instead, we will look at a few selected themes and tendencies that have been the object of discussion in e-lexicography recently.

5.1 Flexible Data Presentation

The use of hyperlinks in a browser leads to different ways of navigating as compared with the two dimensions of a sheet of paper. On the computer screen, you can read from the top left to the lower right corner as on a book page but, in addition, you can also navigate ‘downwards’ by clicking on links that will expand an element on the page or take you to a different page. This has been exploited in e-dictionaries in various ways:

- (1) by having different functionalities on different tabs which the user can shift between
- (2) by letting the user choose between different contents according to a specified profile
- (3) by letting the user expand or unfold certain information types by clicking a button or symbol.

Many lexicographers have seen this as the fulfilment of their dreams and have welcomed the digital possibilities with enthusiasm. Through hyperlinks they

can now present to the user all and only the relevant information needed in a specific look-up situation. Consequently, a number of edictionaries have appeared that make use of the customization possibilities, ranging from fairly simple options (show more, show less) to highly elaborate user profiles (Trap-Jensen 2010; Verlinde 2010) whereas others have marketed the customized content as different dictionaries altogether (Bergenholtz 2011). In a way, it is the lexicographer's dream, but it has turned out to have one serious disadvantage: so far, user studies have not been able to confirm that users take advantage of the possibilities offered to them. On the contrary, evidence suggests that they are not very good at analysing their own needs and the look-up situation they are in (Trap-Jensen 2010, Lorentzen and Theilgaard 2012). Lexicographers will have to respond to this challenge and find new ways of accommodating the users. One possible reaction could involve changing the focus from customization towards the use of adaptive technologies: instead of leaving it to the users to select the appropriate combination of data for a task, the dictionary could do so, adapting in line with the user's previous search behaviour. This is in keeping with the service provided by Amazon and other companies that offer new items to their customers based on what they have bought earlier (cf. Rundell 2012: 23).

5.2 Crowdsourcing and Collaborative Lexicography

While everyone knows Wikipedia, few successful attempts have so far been made at creating dictionaries with content that is entirely user-generated (though see Wiktionary, www.wiktionary.org). This could change, of course, but it seems to be in keeping with a preference among crowdsourcing contributors for niche areas where they are experts. Thus user-driven dictionaries are more likely to be successful if they are directed towards a limited area (such as slang, neologisms, dialects, special subject fields) rather than towards general language vocabulary. Lexicographers should take advantage of this and welcome contributions from users. Most obvious are suggestions for new entries, where users can submit anything from a headword to a full entry proposal with sense divisions, definitions, collocations and authentic examples. User-involvement and interactivity in general are characteristic trends in internet behaviour, which can be incorporated in dictionaries in various forms: on social media, as blogs or forums, RSS feeds, questions and answers, comments and feedback on individual entries, etc. Also entertainment, gamification and dynamic content are features that cannot be dismissed as a mere whim of fashion, especially in learners' dictionaries and other dictionaries aimed at the younger generation of digital natives.

6 Finding the Dictionary – the Future

The digital reality that we live in today is going to change the form and status of the dictionary, no doubt about it. The question is: how will it change? If we think of the analogue products of the not-so-distant past, the dictionary was a very concrete and tangible object: a physical book on a shelf. Faced with a linguistic problem, the user would have to make a deliberate choice and reach out for the dictionary that he or she thought would help solve the problem. This is not so in the digital era. Faced with a similar problem today, nine out of ten people do not turn to their favourite e-dictionary. Instead, they simply ask Google and they don't care if the answer comes from a dictionary, a forum discussion or a newspaper article. One response from lexicographers to this challenge is search engine optimization: make sure your dictionary appears as early as possible on the Google result page. Another reaction is resource integration: provide the answer to the user where the problem occurs. Instead of turning to a completely different site, whether an e-dictionary or Google, the user looks up in the embedded dictionary via a keyboard shortcut (e.g. double click) without leaving the site. This is already common in many e-readers but could be developed further, for instance as part of individual applications and sites or even as part of the computer's operating system. Much more will be said about future dictionaries in Chapter 5.

The challenge for lexicography in digital times is that dictionaries will definitely change their appearance and most likely will lose status and run the risk of drowning in the profusion of other resources with which they compete for user attention. Whether this is viewed as a good or bad thing is more than anything a matter of individual inclination. For the pessimist, it may be a comfort that nothing suggests that the need for lexicographical data is diminishing.

References

- Bergenholtz, H. (2011) Access to and presentation of needs-adapted data in mono-functional internet dictionaries. In: P. A. Fuertes-Olivera and H. Bergenholtz (eds) *e-Lexicography: The Internet, Digital Initiatives and Lexicography*. London and New York: Continuum, 30–53.
- Bergenholtz, H. and Johnsen, M. (2005) Log files as a tool for improving internet dictionaries. *Hermes* 34, 117–41.
- De Schryver, G.-M., Joffe, D., Joffe, P. and Hillewaert, S. (2006) Do dictionary users really look up frequent words? – on the overestimation of the value of corpus-based lexicography. *Lexikos* 16, 67–83.
- Langebek, J. (1740) *Plan for the Organisation of the Dictionary of the Royal Academy*. Here quoted from L. Jacobsen and H. Juul-Jensen (1918) Preface to the *Dictionary of the Danish Language*, Vol. I, Copenhagen: Gyldendal Publishers.
- Lorentzen, H. and Theilgaard, L. (2012) Online dictionaries – how do users find them and what do they do once they have? In: R.V. Fjeld and J. M. Torjusen (eds) *Proceedings*

- of the 15th EURALEX International Congress. Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo, 654–60.
- Lorentzen, H. and Trap-Jensen, L. (2011) There and back again – from dictionary to wordnet to thesaurus and vice versa: how to use and reuse dictionary data in a conceptual dictionary. In: I. Kosem and K. Kosem (eds) *Electronic Lexicography in the 21st Century. New Applications for New Users* (Proceedings of eLex 2011, Bled, 10–12 November 2011). Ljubljana: Trojina, Institute for Applied Slovene Studies, 175–9.
- Molbech, C. (1859) *Molbechs ordbog 1–2*. Copenhagen: Gyldendalske Boghandlings Forlag.
- Pedersen, B. S. et al. (2009) DanNet: the challenge of compiling a wordnet for Danish by reusing a monolingual dictionary. *Language Resources and Evaluation*, 43/3, 269–99.
- Rundell, M. (2012) It works in practice but will it work in theory? The uneasy relationship between lexicography and matters theoretical. In: R. V. Fjeld. and J. M. Torjusen (eds) *Proceedings of the 15th EURALEX International Congress*. Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo, 47–92.
- Trap-Jensen, L. (2010) One, two, many: customization and user profiles in internet dictionaries. In: A. Dykstra and T. Schoonheim (eds) *Proceedings of the XIV Euralex International Congress, Leeuwarden*. Fryske Akademy, 1133–43.
- Verlinde, S. (2010) The base lexicale du français: a multi-purpose lexicographic tool. In: S. Granger and M. Pacquot (eds) *eLexicography in the 21st Century: New Challenges, New Applications* (Proceedings of eLEX 2009). Louvain-la-Neuve: Presses Universitaires de Louvain (Cahiers du CENTAL series), 325–34.