

Towards variable function-dependent sense ordering in future dictionaries

Robert Lew

Adam Mickiewicz University
School of English
Poland
rlew@amu.edu.pl

Towards variable function-dependent sense ordering in future dictionaries

Abstract

The present paper deals with the relationship between sense ordering and the primary function of the dictionary, arguing that the former should reflect the latter. I give suggestions as to how this could be achieved in practice, and which sense ordering would best serve the different functions. Given that lexicography appears to be at an important crossroads, I claim that lexicographers should embrace more enthusiastically the potential afforded by new storage media. With regard to sense ordering I thus propose that we move beyond static ordering in electronic polyfunctional dictionaries of the future, to dynamically adjust the ordering to the currently dominant function. Another future possibility that I explore is that Artificial Intelligence systems in electronic lexicographic products could conceivably monitor individual users' lookup behaviour and thus customize sense ordering on an individual user basis, for example by deprioritizing senses known to have already been acquired by the user, when in the text reception/decoding mode. Finally, I outline a programme of research to provide a more empirical basis for the above.

1. Sense ordering in dictionaries

The issue of sense ordering in dictionaries arises in those entries which have more than a single sense. Such entries have often been referred to in the literature as *polysemous entries*. When talking about *polysemy*, though, we need to be careful to distinguish between polysemy as a linguistic, lexical-semantic concept referring to hypothetical semantic properties of units of language, and a derived sense referring to an aspect of the dictionary (micro)structure. It needs to be borne in mind that a dictionary description does not necessarily match a lexical-semantic structure of a unit of language postulated by linguists. This is so because of the diverging aims of lexicographers and linguists: A good dictionary should be a maximally helpful tool for the target user in a range of problems and consultation situations for which it is being designed. In contrast, a linguistic description of a language is a model of language (whatever the terms 'model' and 'language' may mean within a particular linguistic paradigm). True enough, the two may sometimes converge for some dictionaries: those aiming to faithfully model some part of the lexicon; this, of course, would typically reflect a specific theoretical position regarding the nature of the lexicon. Such convergence, however, should be seen as an exception rather than the rule. All in all, given the above problems, it is probably safer to use the more neutral term *multiple-sense entries* when referring to lexicographic entities, rather than *polysemous entries*.

In multiple-sense dictionary entries and within a linear lexicographic presentation space (see Figure 8 below for an example of the non-linear alternative), a decision must be taken as to the ordering of these multiple senses. The decision as to which sense should be placed first is especially important in view of the available findings from research on dictionary use which indicate that dictionary users tend to stop at the initial sense given, unless they get a very clear signal compelling them to move on (Tono 1984; Lew 2004). In view of this marked preference of users for the first sense(s), it is fair to say that sense-ordering decisions have crucial practical implications for the success of dictionary consultation, and the issue of sense ordering itself deserves more attention from metalexicographers than it has received so far.

Lexicographic decisions with regard to sense ordering typically follow several others, notably the grouping of presentation into discrete senses (witness the continuing debate between 'lumpers' and 'splitters'). Still, some decisions as to grouping and ordering may be interdependent. Even so, in

this paper I will try to focus on sense ordering *per se*, deliberately putting aside any related decisions with respect to sense grouping. I am aware that this is an oversimplification, but it is one that needs to be made at this rather preliminary stage of research into the topic. In the next section I will attempt a brief overview of the ordering strategies noted in the lexicographic literature.

2. Overview of ordering strategies

2.1. Sense ordering strategies

Reference to a number of sense ordering strategies can be found in the lexicographical publications, including theoretical papers and the front matter of dictionaries; however, there are few systematic studies of the actual practice of sense ordering in published dictionaries. Still, what the occasional comments in studies focusing on other aspects as well as some introductions to dictionaries suggest is that quite often no uniform principles are in fact applied across all entries (see e.g. Jackson 2002: 92). The principles themselves, however, have received some attention in the literature, which provides at least the following options with respect to sense ordering:

1. chronological (historical) ordering
2. ordering by frequency (usage)
3. logical ordering
4. textual/pragmatic ordering
5. other local strategies

The criteria assembled under the pseudo-category 5 above are of restricted, local use, in the sense that they tend to be applied to a subset of items and/or for only a few of their senses, and will be discussed last. First, the four global strategies listed above will be taken up in turn.

2.2. Chronological (historical) ordering

The idea behind chronological ordering is to have the senses ordered from oldest attested to most recent ones. The dating of the senses is usually based on the available citation evidence. This evidence, of course, is only as good as it is complete, and quite often it is far from complete, especially for the earlier periods, for which available textual evidence tends to be fragmentary at best. And, such evidence may be lacking altogether for those (quite numerous) languages which happen not to have a large and old enough body of written records. Chronological ordering tends to be the strategy of choice for philological, diachronic historical dictionaries. Even so, many amongst those dictionaries have opted for logical ordering instead, or a mixed system, especially if the ordering is hierarchical rather than flat, with senses being grouped at an intermediate level. Figure 1 below gives an example from the online version of the *Oxford English Dictionary*.

Oxford English Dictionary

Lost for Words? Find Word

Results **summit, n.** SECOND EDITION 1989

Entry

summit, n. **a.** The topmost part, top; the vertex, apex; †the crown (of the head), boss (of a shield), umbo (of a shell).

1470-85 MALORY Arthur v. viii. 174 It clefte his hede fro the somette of his bede. 1471 CAXTON Recuyell (Sommer) 414 Vpon the somette or toppe of the tour, he maad an ymage of copre. Ibid. 615 The maystres had sette on the sommet or toppe of the hede of hector...a vessell. c1477 — Jason 16 Jason smote hym on the sommette of his shelde. c1513 FARNHAM Chron. ii. xxii. (1811) 23 He buylded an Hauen w' a gate ther ouer... In the summit or pynacle wheron was set a vessell of Brasse. 1706 PRIOR Ode to the Queen xxviii, Let Europe say'd the Column high erect... Sublime the Queen shall on the Summit stand. 1784 COWPER Task iii. 536 Golden flow'rs, Blown on the summit of th' apparent fruit. 1828 STARK Elem. Nat. Hist. II. 63 Shell thin...with the summit pointed. 1859 TENNYSON Elaine 482 A wild wave...Green-glimmering toward the summit. 1866 R. TATE Brit. Mollusks iv. 92 The tentacles... bear the eyes at their summits.

b. *Sot.* By an etymologizing alteration of Grew's *SEMET*, used for 'anther'; and hence for 'stigma'. *Obs.*

1731 MILLER Gard. Dict., *Summits* or *Apices* are those Bodies which contain the Prolifick Powder. 1753 Chambers' Cycl. Suppl., App., *Summits* of flowers, the same with the anthers, or top of the stamina. 1796 WATHERING Brit. Plants (ed. 3) I. 5 The Germen, the Summit, and the Anthers, are all that are essentially necessary.

c. *Geom.* A point of a polyhedron where three or more faces meet, forming a solid angle.

1805-17 [see *summit* angle in 4]. 1823 BROOKE Crystalllog. 6 The regular tetrahedron...contained within four equilateral triangular planes. The solid angle at a, is sometimes called its summit.

2. The topmost point or ridge of a mountain or hill. Also, the highest elevation of a road, railway, or canal.

1481 CAXTON Godfrey clxx. 251 Syon is toward the weste, on the sommete or toppe theron stondeth the chirche which is named Syon. 1602 SHAKES. Ham. iii. iii. 18 It is a massie wheele flat on the Sommet of the highest Mount. 1697 DRYDEN Æneis x. 984 He, like a solid Rock by Seas inclos'd...From his proud Summit looking down. c1700 EVELYN Diary (1850) I. 2 Leith Hill, one of the most eminent in England for the prodigious prospect to be seen from its summit. 1736 GRAY Statius ii. 18 Ætna's smoking summit. 1774 COLLIER, Nat. Hist. (1776) I. 155 Some of the summits of the Alps have never yet been visited by man. 1859 DEKKER T. Two Cities I. ii, The last burst carried the mail to the summit of the hill. 1860 TYNDALL Glac. i. iii. 30 To witness the scene from the summit of the pass. 1891 E. ROPER By Track & Trail x. 148 Summit, in railway parlance, means the highest point attained by the line in crossing a mountain.

List by entry List by date Entry map

Single search Advanced search

Print Mail Help Sign out

Subscribe Adam Mickiewicz University

OXFORD UNIVERSITY PRESS

Figure 1. Part of the entry for **summit** in the online version of the *Oxford English Dictionary*

It will be observed in the *Oxford English Dictionary* entry that while the subsenses indicated with lower-case letters **a.**, **b.**, **c.** (as well as the citations under each subsense) can be claimed to be arranged chronologically, the sense groupings marked with roman numerals are no longer so arranged – the likely arrangement in this case is general-to-specific (for which see section 2.4 below).

Apart from historical dictionaries, traditional general monolingual dictionaries may favour historical ordering as well. A case in point would be the *Merriam-Webster's Collegiate Dictionary*, which in the preface (Merriam Webster's Collegiate Dictionary 2004: 20a) explicitly declares its senses to be ordered historically. Interestingly, McCreary (2008: 1293) suggests that a *reverse* historical ordering might be a good policy for this particular dictionary.

2.3. Ordering by frequency

The principle of ordering by frequency (sometimes known as 'ordering by usage') calls for the most frequent sense(s) to be placed first in the entry. This strategy has gained popularity with the growing role of corpora in dictionary-making, in part because corpora can be helpful in assessing the frequencies of different senses and uses. Of course, there are issues with the objectivity of corpus-derived frequencies, since they can be rather sensitive to the composition of the corpus. And, while corpus-trawling software is quite effective at counting forms, it will not (as yet, at least) cluster senses, so a human lexicographer will normally still be needed to decide on which tokens

constitute which senses. This part of the job, being subject to human assessment, will carry the usual baggage of subjectivity with it. And, while in principle one could use linguistic criteria for polysemy such as the ones discussed by Cruse (1986; 1993), they are not always relevant to what is optimal sense structure in practical lexicography, even if time pressures allowed them to be used systematically.

The strategy of frequency-based ordering has been embraced by monolingual English learners' dictionaries, and it has often been claimed without much reflection or evidence that ordering the most frequent senses first gives the user the best chance of quickly locating what they need; but this is yet to be demonstrated, and I return to this issue below. Figure 2 gives an example of an entry for **summit** from an online version of the most recent Fourth Edition of the *Longman Dictionary of Contemporary English*.



Figure 2. The entry for **summit** in the online version of the *Longman Dictionary of Contemporary English*, Fourth Edition

As shown in Figure 2, the 'political meeting' sense is given first, presumably on the evidence of corpus-derived frequency; the promotion of this sense to the privileged first-sense position was probably first effected, and criticized, in the first edition of the *Collins COBUILD English Language Dictionary* of 1987.

Learners' dictionaries, however, are not unanimous in how much weight they attach to corpus frequency data, and they tend to look at other criteria as well. Furthermore, they do not all use the same corpora, so their evidence as to frequency may vary (cf. Bogaards 2008). Thus, an entry for **summit** in another popular English learner's dictionary, the *Oxford Advanced Learner's Dictionary of Current English* in its Seventh Edition reveals the 'political meeting' sense being given at the second, and last, sense of the entry (see Figure 3).

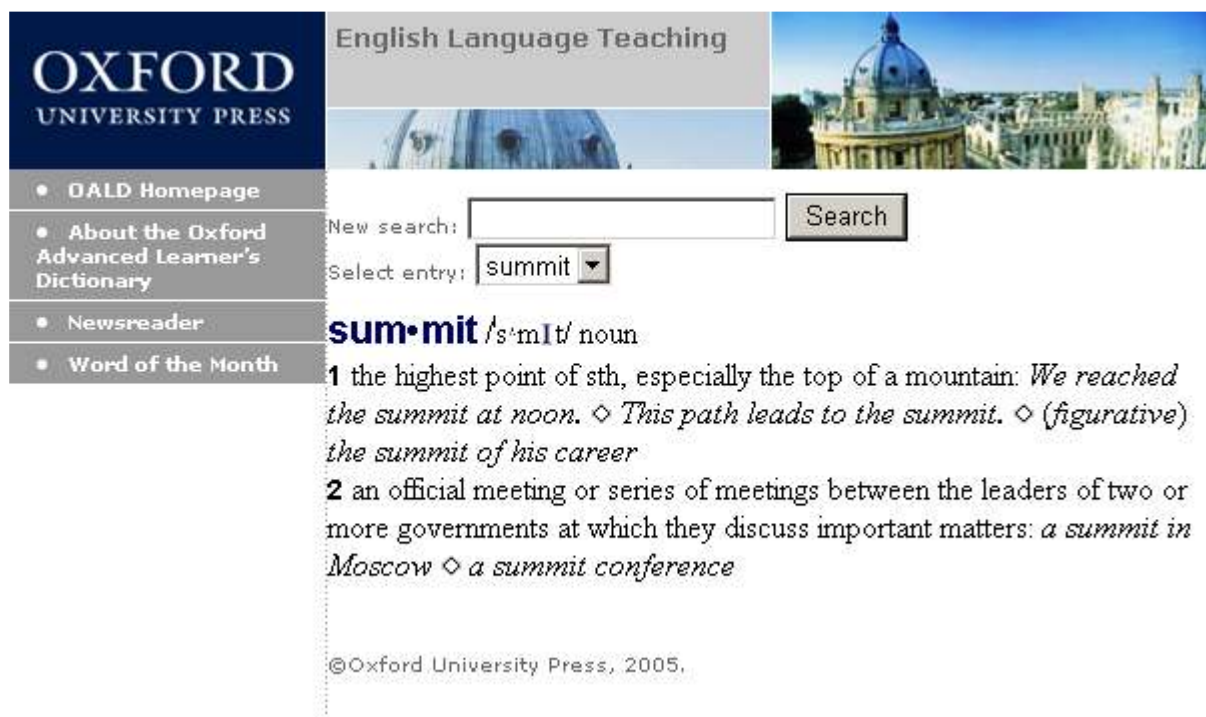


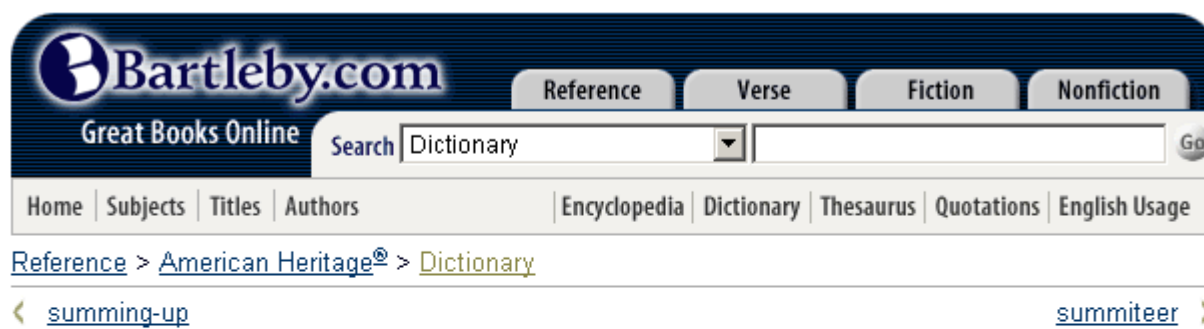
Figure 3. The entry for **summit** in the online version of the *Oxford Advanced Learner's Dictionary of Current English*, Seventh Edition

2.4. Logical ordering

The idea that dictionary senses can be arranged in a 'logical' way presumes that there is such a thing as core or central meaning. Looking through the literature, several dimensions of logical ordering can be found, some of them in overlap, others in contradiction (in specific cases). For Moerdijk (2003: 285) logical ordering implies a hierarchical (as opposed to flat) sense structure. Logical ordering includes at least the following aspects:

- from central/core to peripheral
- from general to specific
- from concrete (spatial, temporal) to abstract
- from literal to metaphorical
- from original to derived

A few additional comments will be offered with reference to the above dimensions. The core-to-peripheral principle is an important one in *The New Oxford Dictionary of English* (Pearsall 1998), which attempts to identify a core sense and list it first (followed by derivative subsenses). The general-to-specific dimension is related to William Frawley's principle of *semantic inclusion* (Frawley 1989: 235). A good example of an entry arranged in this way may be the entry for **summit** in the *The American Heritage Dictionary of the English Language*, Fourth Edition, given in Figure 4.



[CONTENTS](#) · [INDEX](#) · [ILLUSTRATIONS](#) · [BIBLIOGRAPHIC RECORD](#)

The American Heritage® Dictionary of the English Language: Fourth Edition. 2000.

summit

SYLLABICATION: sum-mit

PRONUNCIATION:  sūm'it

NOUN: **1.** The highest point or part; the top. **2.** The highest level or degree that can be attained. **3a.** The highest level, as of government officials. **b.** A conference or meeting of high-level leaders, usually called to shape a program of action.

VERB: Inflected forms: **sum-mit-ed**, **sum-mit-ing**, **sum-mits**

TRANSITIVE VERB: To climb to the summit of (a mountain).

INTRANSITIVE VERB: To climb to the summit.

ETYMOLOGY: Middle English *somet*, from Old French *sommette*, diminutive of *som*, top, from Latin *summum*, from neuter of *summus*, highest. See [uper](#) in Appendix I.

Figure 4. The entry for **summit** in *The American Heritage Dictionary of the English Language* (Fourth Edition, online version)

Interestingly, James Murray (1888), the editor of the Oxford English Dictionary, argues in the front matter to OED that historical ordering actually *is* logical. Stockwell and Minkova (2001: 188), on the other hand, point to a parallel between logical ordering and frequency ordering. Such opinions suggest that there is a considerable degree of fuzziness in the category of logical ordering. In turn, Kipfer (1984: 103-104) seems to distinguish logical ordering from ‘psychologically-meaningful’ ordering, giving the latter a separate section in her paper and recommending it as superior, but the actual differences are not at all clear to me from Kipfer’s work.

2.5. Textual/pragmatic ordering

This ordering strategy was proposed within the *dictionary as text view* (Frawley 1989). It postulates that senses are best ordered in a way that creates a progression from given information to new information, so that a dictionary text abides by the typical structure of texts in general.

2.6. Other local strategies

This last category is a waste basket for such strategies that are only applicable to a subset of dictionary entries and/or senses. One of those is syntactic ordering, which dictates that senses be arranged based on their specific syntactic properties. This strategy would not normally be applied across the lexicon, but it could be used, on occasion, to order verb entries according to their complementation patterns, or arrange entries for grammatical words based on specific syntactic patterns of use. Other local strategies noted in the literature would be putting the senses applying to people in front of those applying to things; and neutral-connotation senses before marked (positive and negative) ones (Fraser 2008). This last criterion might be seen as a special case of logical ordering if we interpret the general-specific contrast as being instantiated in the opposition between unmarked and marked. A related opposition can also be recognized in the occasional tendency to push towards the end of an entry senses with restricted usage, such as those marked for currency, register and domain (e.g. obsolete, slang, or technical senses).

2.7. Relevance of sense ordering strategies to bilingual dictionaries

Just like for the majority of theoretical lexicographic writing, which is dominated by monolingual dictionaries, the above strategies have mostly been described with monolingual dictionaries in mind. When it comes to bilingual dictionaries, it seems that the above strategies might still be relevant as long as the sense structure is based on that of the source language of the dictionary. This is frequently the case when bilingual dictionaries are created by translating monolingual dictionary frameworks, not an uncommon scenario.

On the other hand, the strategies outlined above may lose or change their significance in those bilingual dictionaries whose microstructural organization is built around target language equivalents (equivalence structure, Manley et al. 1988; target-language structure, Adamska-Sałaciak 2006: 66). I shall have more to say on sense ordering in such bilingual dictionaries in section 4.3 below.

3. The quest for optimal ordering strategies

3.1. Optimal sense ordering is relative

The discussion so far prompts the question of whether a search for the best ordering strategy makes any sense at all. I believe that an expectation that a single strategy might be optimal in all circumstances is naïve and unrealistic. There are at least two important reasons for this.

First, there is no good reason to assume *a priori* that the same sense ordering will work best for all uses of a dictionary; rather, the decision as to how senses should best be ordered ought to reflect (just as many other important lexicographic decisions!) the design function(s) of the dictionary. In fact, Kipfer (1984) already appears to be steering towards a similar conclusion in some passages, though she backtracks in the conclusion, proposing instead uniformity across dictionaries from different publishers in the interest of ‘serving the user best’ (Kipfer 1984: 108). In this contribution I would like to develop the idea of variable and even dynamic sense ordering and give it some substance.

Second, experienced lexicographers are aware that, the benefits of a uniform dictionary stylesheet notwithstanding, some items are best treated in an individual way, as different lexicographic items may call for special lexicographic treatment. This sentiment is elegantly captured by Patrick Hanks’ (2006, personal communication¹) principle of *horses for courses*. So, the best sense-ordering strategy may actually be ... not to have a strategy, or, more precisely, to aim for item-specific treatment.

Optimal sense ordering may depend on the function of the dictionary (for a background on lexicographic functions see e.g. Tarp 1992, 1998; Bergenholtz and Tarp 2002, 2003, 2004, 2005; Tarp 2008). For example, historical ordering in philological dictionaries may serve scholars in

language research, it is also a sensible choice in etymological dictionaries, since entries in such dictionaries are primarily about the development of words across time – see for example the entry for **log** in the *Online Etymology Dictionary* at <http://www.etymonline.com/index.php?term=log>, reproduced in Figure 5.

The screenshot shows the Online Etymology Dictionary interface. At the top is a logo of a winged figure holding a shield, with the text "ONLINE ETYMOLOGY DICTIONARY" below it. A search bar contains the word "log", and a "Search Mode" dropdown is set to "All". Below the search bar is a horizontal menu of letters from A to Z. The entry for "log" is displayed below the menu. It includes two sections: "log (n.)" and "log (v.)". The "log (n.)" section provides a detailed etymology, starting with "1398, of unknown origin. O.N. had *lag* 'felled tree' (from stem of *leggja* 'to lie'), but on phonological grounds etymologists deny that this is the root of Eng. *log*. Instead, they suggest an independent formation meant to 'express the notion of something massive by a word of appropriate sound.' Logging 'act of cutting timber' is from 1706. *Lagiam* 'congestion of logs on a river' is from 1885; in the figurative sense it is from 1890. *Logrolling* in the legislative vote-trading sense first recorded 1823, from the notion of neighbors on the frontier helping one another with the heavy work of clearing land and building cabins (as in phrase *you roll my log and I'll roll yours*). *Log cabin* in Amer.Eng. has been a figure of the honest pioneer since the 1840 presidential campaign of William Henry Harrison. The "log (v.)" section provides a brief etymology: "to enter into a log book," 1823, from *logbook* "daily record of a ship's speed, progress, etc." (1679), which is so called because wooden floats were used to measure a ship's speed. To *log in* in the computing sense is attested from 1963. At the bottom of the page, there is a footer with navigation links: "Introduction and abbreviations", "Who did this?", "Sponsored Words", and "Links". On the right side of the footer, there is copyright information: "© November 2001 Douglas Harper", "Logo design by LogoBee.com", and "Page design and coding by Dan McCormack".

Figure 5. The entry for **log** in the *Online Etymology Dictionary* (<http://www.etymonline.com/index.php?term=log>)

In turn, the very popular English monolingual learners' dictionaries in their majority tend to list the most common senses first. But, learners' dictionaries (and not just them!) are typically polyfunctional: that is, they have more than one clearly dominant function.

Sense ordering – and its relationship to lexicographic functions – take on new significance in the context of modern electronic media that lexicography is rapidly moving to. One exciting consequence of modern media is dynamic presentation of lexicographic data. Given that lexicography appears to be at an important crossroads, I would suggest that lexicographers should embrace more enthusiastically the potential afforded by new storage media, including dynamic presentation. Focusing on sense ordering, perhaps it could be made to suit the *currently active* function in dynamic presentation interfaces of (future) electronic dictionaries. It is this idea that I turn to in the next section, where different functions are presented as subsections.

4. Function-dependent sense ordering: preliminary proposals

4.1. Reception (text comprehension)

We know from numerous user studies that text reception is the most common use dictionaries are put to. Let us take the case of English Monolingual Learners' Dictionaries (MLD's). Following Hornby's (1965) recommendation, these dictionaries tend to follow the 'most frequent senses first' principle. Hornby's argument goes as follows:

another problem is that of the order in which semantic varieties are to be placed. For those to whom English is a foreign language, those who are learning (not necessarily ‘studying’) English for present-day needs, the order in which semantic varieties are entered should perhaps be based on frequency rather than on historical principles. Such users of a dictionary are more likely to meet, and to need to use, words in their current senses. These are the senses, therefore, which should be entered first. (Hornby 1965: 106)

The reason I have quoted Hornby extensively above is that his suggestion to use frequency ordering appears to have been taken to heart by many lexicographers, even though it is couched in somewhat tentative terms (*‘should perhaps’*). It should be noted that Hornby’s interests lay probably more with productive dictionary use (*‘need to use’*), which is discussed in the following section. Within the context of text comprehension, however, Hornby’s proposal is controversial, especially in view of the rather undisputed fact that the major users of MLD’s are relatively advanced learners. (And for a very good reason: there is a limit to which one can simplify *foreign language* definitions, so they *must* remain largely incomprehensible to lower-proficiency learners.) Now, my point is that such relatively advanced learners are quite likely to be fairly familiar with the most frequent senses of highly polysemous (and thus also frequent, by Zipf’s law) lexical items.

In view of the above, if editors of MLD’s insist on being guided by frequency, there are other options to consider, such as listing *medium-frequency* items first. Other than frequency ordering, the *general before specific* ordering seems an interesting possibility for text comprehension purposes, as it can perhaps promote partial comprehension after exposure to initial (=general) senses.

4.2. Text production; text revision

Unlike for text reception, non-native speakers using monolingual dictionaries in text production and revision would probably be well served by dictionary entries ordered by descending frequency, and in this I agree with Hornby. The counterargument presented in the previous section about most frequent senses being already known loses much of its punch because what is often needed in production and revision are fine points of collocation and grammar, and native-like *active* mastery of words is reached far less easily than mere passive comprehension. Another independent point speaking in favour of frequency ordering is that it would be beneficial to assist learners in using more often just those senses and uses that are most frequent in native texts. After all, item overuse and underuse are problems that persist well into very advanced stages of foreign language proficiency.

4.3. Text production and revision, translation into L2 with a bilingual dictionary

With regard to bilingual L1->L2 dictionaries for foreign language production, if one is willing to agree with the sensible assumption that a person’s native language provides the most efficient path of access to foreign language forms (Piotrowski 1994: 78; Lew 2004), descending translational frequency could be used as a functional criterion for sense ordering, with the most frequent (cross-textual) equivalent(s) being listed first in the bilingual entry, followed by less frequent ones. A similar argumentation applies to L1->L2 translation, of which the production of a natural-sounding foreign text is characteristically the most challenging stage. Text revision involves a varying degree of new text production, and in this respect it is partially similar to text production proper, but since the starting point is the target-language text, a reference work accessible through forms of that language would be comparatively more important. Here too, though, the ordering of senses by descending frequency seems sensible for reasons outlined above.

Of course, it would be naïve to believe that all possible target language equivalents can be found, or placed, in dictionary entries. This is in principle impossible, since lexical choice is not just a local

issue, and depends on broader cotext and context which for the moment cannot be comprehensively accounted for in dictionary entries, at least in the current understanding of what a dictionary is (Piotrowski 2007). But I think we can make a fairly good start in the right direction by providing just those equivalents to translators that turn out to be most frequent translations in equivalent texts. There will be frequent occasions when such equivalents will work just fine, and at some other times the equivalents supplied may provide a useful departure in a search for other options. Once again: the above comments on bilingual dictionaries apply to text production, revision, and translation. In text comprehension the best equivalents may well be different: cognitive equivalents rather than translation or functional equivalents (for the distinction see e.g. Piotrowski 1994; Adamska-Sałaciak 2006: 101-106).

4.4. Studying and browsing

Studying and browsing might be described as serving cognitive functions (compare cognitive situations, be it communication-related or not, Tarp 2007). An important desirable outcome of studying and browsing is that they should result in lexical acquisition and/or knowledge acquisition. Here, logical (psychologically-meaningful?) ordering of senses suggests itself as a potentially promising option, since this kind of ordering could be conducive to lexical and knowledge acquisition.

5. Towards an empirical basis for sense ordering decisions

The above proposals for the best sense ordering solutions depending on lexicographic functions have to be viewed as very tentative. They are based on theoretical rationale, but, for the most part, lack any empirical basis; they are thus, at this time, little more than informed guesses. Before such recommendations can be given with any assurance, we need to have a rather more substantial foundation on which to build specific lexicographic guidelines and dictionary-making practice. In what follows, I propose some ways to pursue such an empirical basis.

5.1. Text comprehension

When it comes to reading text comprehension (dictionaries are only very rarely used in *listening* comprehension), we could propose user studies of lookup patterns in text comprehension recording the dictionary senses wanted. Such studies might ultimately take into account the role of two types of context:

- the textual context of the text being read
- the context of the dictionary text itself (e.g. neighbouring senses, sense indicators, various other aspects of entry organization)

It seems that both these types of context might have an impact on the process of sense selection by the dictionary user. Of course, ‘the dictionary user’ is a carrier of a set of complex, stratified variables, and research design should ultimately take account of this.

It might also be possible to capitalize here on the findings of existing research into vocabulary growth in language learners (both young and mature), as well as native speaking children, with implications for the relevant dictionary types. Data on vocabulary growth could fruitfully inform the sense ordering decisions in future dictionaries.

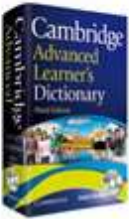
In a recent paper, de Schryver *et al.* (2006) discuss evidence from online dictionary logs suggesting that, contrary to common claims, rare words are looked up just as often as frequent words. Such log-based studies are extremely interesting (see also Bergenholtz and Johnson 2005, 2007). In de Schryver *et al.*’s study logged records are available for the look-up queries which on the whole correspond to lemma signs; however, for the purpose of sense-ordering we would need finer-

grained information. How, then, can we find out about the frequency of looking up specific senses? The good news is that, in fact, some existing electronic dictionary interfaces can potentially be used to trace such usage data without much modification.

As an example, consider a menu driven interface of the online version of the *Cambridge Advanced Learners' Dictionary* (Figure 6).

CAMBRIDGE | Dictionaries Online

Cambridge Advanced Learner's Dictionary

 [Buy me!](#)

[Look it up](#)

Select another dictionary.. ▾

We publish dictionaries for people learning English all over the world.
[Find out more...](#)

Results

relate was found in the [Cambridge Advanced Learner's Dictionary](#) at the entries listed below.

- [relate \(CONNECT\)](#)
- [relate \(TELL\)](#)
- [relate to sb/sth \(CONNECT\)](#)
- [relate to sb \(UNDERSTAND\)](#)
- [relate to sth](#)

Figure 6. The initial entry for **relate** in the online version of the *Cambridge Advanced Learners' Dictionary*, Second Edition

It seems that an interface of this form could readily be used to log data on looking up specific senses. In the online version of the *Cambridge Advanced Learners' Dictionary*, to go to a specific sense of a multiple-sense entry, the user must click on the sense of choice in an entry menu. Only then is the appropriate sense of the actual full entry presented on screen. These sense-selecting clicks of the user could be logged and then analyzed. It should, though, be borne in mind that the degree to which a sense menu like that in Figure 6 is affected by the users' preference for early senses is likely smaller than a complete entry, because extremely short items in a list are easier to scan than full entries.

Another popular English learner's dictionary, the *Longman Dictionary of Contemporary English* (Fourth Edition) in its online version also affords an opportunity for the logging of sense lookups. In Figure 7 the entry for the verb **run** is reproduced from this dictionary.

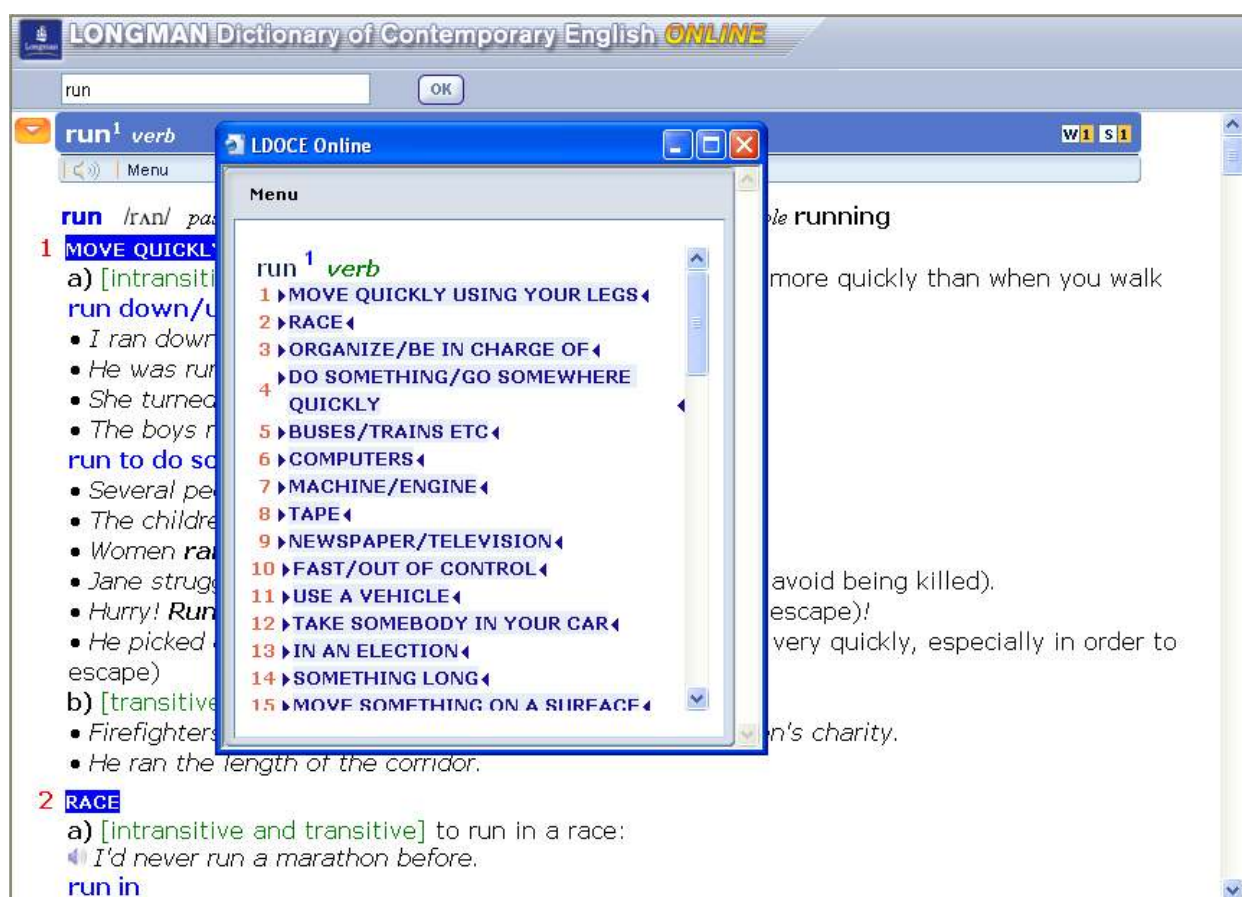


Figure 7. The entry for the verb **run** in the online version of the *Longman Dictionary of Contemporary English*, Fourth Edition

The mechanism here is different from that in the *Cambridge Advanced Learners' Dictionary*. The menu opens in a smaller separate pop-up window at the user's request, by clicking on the word *Menu* which appears against a blue background immediately below the headword, next to the loudspeaker symbol which plays back the audio recording of the word's pronunciation. When a line on the menu window is clicked, the main dictionary window in the background scrolls to the sense requested while the menu window remains in the foreground. From a technical point of view, the cross-reference is implemented through a client-side java script function, and thus does not send a unique URL request to the server. Still, the clicks are potentially loggable.

An even more advanced idea is that perhaps empirical data from user behaviour could feed into sense ordering directly, that is, not through the cycle of user research informing new dictionary design, but more directly via an artificial intelligence mechanism built into the electronic reference work itself.² Still, as long as there will be other reference works being designed, electronic and printed, explicit knowledge about user behaviour and its systematic analysis, including comparisons across different classes of users and uses, will also be needed.

5.2. Text production

As described above, based on the assumption that one's native language is the most efficient access path to target-language items, bilingual dictionaries might use translation frequency for a language pair as a guide to sense ordering, especially in dictionaries adopting the target-structure approach to the bilingual entry. But how can we know which lexical units are frequent equivalents across texts in two different languages? One type of answer can come from modern corpus-based linguistics and

language engineering. Specifically, parallel corpora with sophisticated corpus-processing software can provide online information on the relative frequencies of different equivalents. This methodology can be traced back to Reinhard Hartmann's early ideas (see e.g. Hartmann 1985), and one recent practical implementation for the language pair Polish-English which includes the creation of a large, carefully collected parallel corpus and automatic extraction and ranking of equivalents is Krynicki (2006).

5.3. *Cognitive functions*

Cognitive functions have been claimed to arise in the so-called cognitive situations; Tarp (2007) lists the following cognitive situations:

1. while reading: the sudden need to get additional encyclopaedic knowledge in order to understand the text;
2. while writing: the need to know more about a given topic in order to finish a text;
3. during discussions with other people: the need to clarify a specific issue;
4. during processes in the subconscious: the sudden desire to examine something;
5. during dictionary consultation: the desire to know more about a specific topic;
6. during preparation for specialised translation and interpretation tasks: the need to learn more about the subject field in question;
7. in relation to a teaching programme or a course of study: the need to know more about a specific subject field;
8. (...)

Tarp (2007: 175)

Naturally, this list is still speculative and open-ended, as the original ellipsis at point 8 indicates. Specific cognitive situations need to be described by systematically studying dictionary user behaviour in a variety of contexts.

Psycholinguistic research on vocabulary acquisition might also provide a fruitful basis for fleshing out the details. For instance, online experiments are conceivable that would measure the priming effect of previous dictionary senses. Such research could potentially throw light on what it really takes for an arrangement of senses to be 'psychologically meaningful' (Kipfer 1984: 103-104).

6. Customization

The electronic age in dictionaries brings with it great potential for customization (see e.g. de Schryver 2003). Although, as far as I am aware, there has been no mention in the literature of customization of sense ordering specifically, such customization would make perfect sense in view of the preceding discussion of the relationship between optimal sense ordering and function. It is true that traditional paper dictionaries are often polyfunctional, if only because there is a practical limit on how many separate paper dictionaries can be produced: for example, there are very few (if any) complementary pairs of dictionaries for production and reception for the same language pair and direction, even though theoretical awareness of such a possibility has existed for some decades now (compare Iannucci 1957; Al-Kasimi 1984). Electronic reference works tend to be even more universal, but being universal as a product is no longer incompatible with tool specialization, because in an electronic product the presentation and interface can be dynamically adjusted. In the present context, customization in sense ordering can be relevant to:

1. function specificity
2. item specificity
3. domain specificity

These three aspects will be taken up briefly below.

6.1. Function specificity

In polyfunctional electronic dictionaries, sense ordering could conceivably be adjusted to suit the currently active function, which in a broader sense could include aspects of the global dictionary use situation (context of consultation), the more local specific need, and the user profile.

How would an electronic dictionary know which function to activate? There are a number of options open here. There might be a user-selectable option for the different pre-programmed functions. Or, the dictionary might take a more active role and interact with the user to determine the current function, for example, by presenting the user with specific prompts requiring feedback. Finally, in the long term, the dictionary might be programmed to adjust by monitoring a number of factors in the environment, such as what applications are running in the foreground (web browser, PDF viewer – text reception more likely; word processor – text production more likely) and peripheral device activity (mouse activity – text reception more likely; heavy keyboard activity – text production more likely). In this advanced mode of customization the functions would be discovered and switched back and forth in the background, as if behind the user's back. Still, a user-controlled setting to activate/deactivate this kind of Artificial Intelligence function switching would probably be a welcome feature.

6.2. Item specificity

Item specificity in this context refers to modifying the order of sense presentation according to specific dictionary headwords. For example, while in text-comprehension mode, the electronic dictionary (or some intelligent tutoring system of which the dictionary would form one component) might monitor its users' lexical acquisition process and place first those senses which have not yet been acquired by a particular user. In contrast, senses known to have been fully acquired could be placed towards the end, or perhaps even filtered out from the display altogether.

6.3. Domain specificity

Domain specificity (see also Rundell, this volume) would refer to the adjustment of sense ordering according to the subject domain of the text that a user is working on. For instance, if someone is reading a text on astronomy, senses relevant to this domain could be prioritized. Further, in some situations it may be possible to pinpoint the single relevant sense, if a context is available (as in on-line text reception) and a dictionary is able to perform intelligent sense disambiguation; or if the desired sense is specifically indicated by the advanced user (as in text production).

7. Final comments

In the world of electronic dictionaries, not all presentation interfaces have to be linear: one example of non-linear presentation using spatial arrangement of senses is the *Visuwords* dictionary illustrated in Figure 8.

Notes

1. Patrick Hanks proposed this proverb as an alternative title to Michael Rundell's presentation 'More than one way to skin a cat: Why full-sentence definitions have not been universally adopted' at the 2006 Euralex Congress
2. I am indebted to Włodzimierz Sobkowiak for making this suggestion

References

A. Dictionaries

- Hanks, Patrick, Pearsall, Judy, et al.** 1998. *The New Oxford Dictionary of English*. Oxford: Oxford University Press.
- Merriam-Webster's Collegiate Dictionary*. 11th ed. Springfield: Merriam-Webster Incorporated. 2004.
- Murray, James A. H., Simpson, John A. and Weiner, Edmund S. C. (eds.)**. 1989. *Oxford English Dictionary*. (Second Edition.) Oxford: Clarendon Press.
- Oxford English Dictionary Online*. <http://www.oed.com/>
- Pickett, Joseph**. 2000. *The American Heritage Dictionary of the English Language*. (Fourth Edition.) Boston: Houghton Mifflin.
- Sinclair, John, Hanks, Patrick, et al.** 1987. *Collins COBUILD English Language Dictionary*. London - Glasgow: Collins.
- Summers, Della (ed.)** 2003. *Longman Dictionary of Contemporary English*, (Fourth Edition.) Harlow: Longman. Online version at <http://pewebdic2.cw.idm.fr/>
- Visuwords™ online graphical dictionary*. <http://www.visuwords.com/> [based on the WordNet database]
- Walter, Elizabeth (ed.)** 2005. *Cambridge Advanced Learners' Dictionary*. (Second Edition.) Cambridge: Cambridge University Press. Online version at <http://dictionary.cambridge.org/>
- Wehmeier, Sally (ed.)** 2005. *Oxford Advanced Learner's Dictionary of Current English*. (Seventh Edition.) Oxford: Oxford University Press. Online version at <http://www.oup.com/elt/catalogue/teachersites/oald7/lookup?cc=pl>

B. Other literature

- Adamska-Salaciak, Arleta**. 2006. *Meaning and the Bilingual Dictionary: The Case of English and Polish*. Frankfurt am Main: Peter Lang.
- Al-Kasimi, Ali M.** 1984. 'The Interlingual/Translation Dictionary' in Hartmann, R. R. K. (ed.), *Lexicography: Principles and Practice*. London: Academic Press, 153-162.
- Bergenholtz, Henning and Johnson, Mia**. 2005. 'Log Files as a Tool for Improving Internet Dictionaries.' *Hermes* 34: 117-141.
- Bergenholtz, Henning and Johnson, Mia**. 2007. 'Log Files Can and Should Be Prepared for a Functionalistic Approach.' *Lexikos* 17: 1-20.
- Bergenholtz, Henning and Tarp, Sven**. 2002. 'Die moderne lexikographische Funktionslehre. Diskussionsbeitrag zu neuen und alten Paradigmen, die Wörterbücher als Gebrauchsgegenstände verstehen.' *Lexicographica* 18: 253-263.

- Bergenholtz, Henning and Tarp, Sven.** 2003. 'Two Opposing Theories: On H.E. Wiegand's Recent Discovery of Lexicographic Functions.' *Hermes* 31: 171-196.
- Bergenholtz, Henning and Tarp, Sven.** 2004. 'The Concept of 'Dictionary Usage'' in Dollerup, C. (ed.), *Worlds of Words. A Tribute to Arne Zettersten* [Special Issue of Nordic Journal of English Studies 3.1.]. Copenhagen: University of Copenhagen, 23-36.
- Bergenholtz, Henning and Tarp, Sven.** 2005. 'Wörterbuchfunktionen' in Barz, I., Bergenholtz, H. and Korhonen, J. (eds.), *Schreiben, Verstehen, Übersetzen, Lernen. Zu ein- und zweisprachigen Wörterbüchern mit Deutsch*. Frankfurt am Main: Peter Lang, 11-25.
- Bogaards, Paul.** 2008. 'Frequency in Learners' Dictionaries' in Bernal, E. and DeCesaris, J. (eds.), *Proceedings of the XIII Euralex International Congress*. Barcelona: Universitat Pompeu Fabra, 1231-1236.
- Cruse, D. Allan.** 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.
- Cruse, D. Allan.** 1993. 'On Polylexy.' *Dictionaries: The Journal of the Dictionary Society of North America* 14: 88-96.
- de Schryver, Gilles-Maurice.** 2003. 'Lexicographers' Dreams in the Electronic-Dictionary Age.' *International Journal of Lexicography* 16.2: 143-199.
- de Schryver, Gilles-Maurice, Joffe, David, Joffe, Pitta and Hillewaert, Sarah.** 2006. 'Do Dictionary Users Really Look up Frequent Words? – on the Overestimation of the Value of Corpus-Based Lexicography.' *Lexikos* 16: 67-83.
- Fraser, Bruce L.** 2008. 'Beyond Definition: Organising Semantic Information in Bilingual Dictionaries.' *International Journal of Lexicography* 21.1: 69-93.
- Frawley, William.** 1989. 'The Dictionary as Text.' *International Journal of Lexicography* 2.3: 231-248.
- Hartmann, Reinhard Rudolf Karl.** 1985. 'Contrastive Text Analysis and the Search for Equivalence in the Bilingual Dictionary' in Hyldegaard-Jensen, K. and Zettersten, A. (eds.), *Symposium on Lexicography II. Proceedings of the Second International Symposium on Lexicography, May 1984, at the University of Copenhagen, Lexicographica Series Maior 5*. Tübingen: Niemeyer, 121-132.
- Hornby, A. Ibert Sidney.** 1965. 'Some Problems in Lexicography.' *English Language Teaching* 19.104-110.
- Iannucci, James E.** 1957. 'Meaning Discrimination in Bilingual Dictionaries: A New Lexicographical Technique.' *Modern Language Journal* 41.6: 272-281.
- Jackson, Howard.** 2002. *Lexicography: An Introduction*. London/New York: Routledge.
- Kipfer, Barbara Ann.** 1984. 'Methods of Ordering Senses within Entries' in Hartmann, R. R. K. (ed.), *Lexeter '83 Proceedings: Papers from International Conference on Lexicography at Exeter, 9- 12 Sept. 1983, Lexicographica Series Maior 1*. Tübingen: Niemeyer, 101-108.
- Krynicky, Grzegorz.** 2006. *Compilation, Annotation and Alignment of a Polish-English Parallel Corpus*. PhD Dissertation. School of English, Adam Mickiewicz University, Poznań.
- Lew, Robert.** 2004. *Which Dictionary for Whom? Receptive Use of Bilingual, Monolingual and Semi-Bilingual Dictionaries by Polish Learners of English*. Poznań: Motivex.
- Manley, James, Jacobsen, Jane R. and Pedersen, Viggo Hjørnager.** 1988. 'Telling Lies Efficiently: Terminology and the Microstructure in the Bilingual Dictionary' in Hyldegaard-Jensen, K. and Zettersten, A. (eds.), *Symposium on Lexicography III*. Tübingen: Max Niemeyer, 281-301.
- McCreary, Don R.** 2008. 'Looking up "Hard Words" for a Production Test: A Comparative Study of the NOAD, MEDAL, AHD, and MW Collegiate Dictionaries' in Bernal, E. and DeCesaris, J. (eds.), *Proceedings of the XIII Euralex International Congress*. Barcelona: Universitat Pompeu Fabra, 1287-1293.

- Moerdijk, Fons.** 2003. 'The Codification of Semantic Information' in van Sterkenburg, P. (ed.), *A Practical Guide to Lexicography, Terminology and Lexicography Research and Practice* 6. Amsterdam: John Benjamins, 273-296.
- Murray, James A. H.** 1888. 'General Explanations' in Murray, J. A. H. (ed.), *Oxford English Dictionary*. Oxford: Oxford University Press, xxiv–xxxiv.
- Piotrowski, Tadeusz.** 1994. *Problems in Bilingual Lexicography*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Piotrowski, Tadeusz.** 2007. 'Lexicography, Semantics and Pragmatics' in Fabiszak, M. (ed.), *Language and Meaning. Cognitive and Functional Perspectives*. Frankfurt am Main: Peter Lang, 225-233.
- Stockwell, Robert and Minkova, Donka.** 2001. *English Words: History and Structure*. New York: Cambridge University Press.
- Tarp, Sven.** 1992. *Prolegomena til teknisk ordbog*. PhD Thesis. Spansk Institut, Aarhus School of Business, Aarhus.
- Tarp, Sven.** 1998. 'Leksikografien på egne ben. Fordelingsstrukturer og byggedele i et brugerorienteret perspektiv.' *Hermes, Journal of Linguistics* 21: 121-137.
- Tarp, Sven.** 2007. 'Lexicography in the Information Age.' *Lexikos* 17: 170-179.
- Tarp, Sven.** 2008. *Lexicography in the Borderland between Knowledge and Non-Knowledge: General Lexicographical Theory with Particular Focus on Learner's Lexicography*. (Lexicographica Series Maior 134.). Tübingen: Niemeyer.
- Tono, Yukio.** 1984. *On the Dictionary User's Reference Skills*. B.Ed. Thesis, Tokyo Gakugei University, Tokyo.