

---

# An Analysis of *The Oxford Guide to Practical Lexicography* (Atkins and Rundell 2008)

Gilles-Maurice de Schryver, *Department of African Languages and Cultures, Ghent University, Ghent, Belgium; Xhosa Department, University of the Western Cape, Bellville, Republic of South Africa; and TshwaneDJe HLT, Pretoria, Republic of South Africa (gillesmaurice.deschryver@UGent.be)*

---

**Abstract:** Since at least a decade ago, the lexicographic community at large has been demanding that a modern textbook be designed — one that would place corpora at the centre of the lexicographic enterprise. Written by two of the most respected practising lexicographers, this book has finally arrived, and delivers on very many levels. This review article presents a critical analysis of its features.

**Keywords:** LEXICOGRAPHY, LEARNERS' DICTIONARY, MONOLINGUAL, BILINGUAL, CORPUS, FRAME SEMANTICS, ENGLISH, FRENCH, TEXTBOOK

**Samenvatting:** Een analyse van *The Oxford Guide to Practical Lexicography* (Atkins en Rundell 2008). Al minstens tien jaar lang eist de volledige lexicografische gemeenschap dat een modern tekstboek zou worden ontworpen — één dat corpora in het centrum van de lexicografische belangstelling zou plaatsen. Geschreven door twee van de meest gerespecteerde praktiserende lexicografen, is dit boek er nu eindelijk, en het ontgoochelt niet. Dit recensie-artikel analyseert de kenmerken ervan kritisch.

**Slutelwoorden:** LEXICOGRAFIE, LEERWOORDENBOEK, VERKLAREND (MONOLINGUAAL), VERTALEND (BILINGUAAL), CORPUS, FRAME SEMANTICS, ENGELS, FRANS, TEKSTBOEK

What matters — and this is critical — is not the writer's intention but the reader's interpretation. (Atkins and Rundell 2008: 439)

## 1. Striking gold

Lexicographers worldwide rejoice — the definite textbook on how to make a dictionary from a corpus has arrived! It is called *The Oxford Guide to Practical Lexicography*, for short OGPL, and was written by B.T. Sue Atkins and Michael Rundell. During the book launch on 17 July 2008, at the Thirteenth EURALEX International Congress in Barcelona, John Davey, Linguistics Editor at OUP UK, simply stated that he had 'struck gold'. He could not have summarized it any better.

In the first edition of the *Macmillan English Dictionary for Advanced Learners*, MED-1 (2002), 'strike gold' is sense 13a under **strike**<sup>1</sup> *verb*. In the second edition, MED-2 (2007), 'strike gold' is listed in a separate phrase block under the same headword. In the latest *Collins-Robert French Dictionary*, CRFD-8 (2006), one reads under sense 2 *transitive verb*, subsense c: '**to strike gold** (figurative use): trouver le filon\*' (where the asterisk marks that the French equivalent, but not the English, is informal). One and the same expression, three (very) different treatments. Is one to conclude — right away — that lexicography is not an exact science, that there are no theories that govern it? Not at all: About every page of OGPL reconfirms both the highly scientific nature of the lexicographic enterprise and the need for strong theoretical underpinnings. Theories evolve, users become ever more demanding, the types of dictionaries spiral outward. Variation and continual improvement, therefore, are inherent. Although reference works have been produced for thousands of years, OGPL proves that the field of lexicography is more alive than ever.

## 2. Importing, exporting and reimporting gold

With reference to the line of dictionaries cited in Section 1, Sue Atkins was the General Editor of the first *Collins-Robert English-French Dictionary* (Atkins and Robert 1978). Michael Rundell has been the Editor-in-Chief of the Macmillan family of learners' dictionaries since their inception. Both were involved in numerous other dictionary projects, beginning in 1966 for Sue, in 1980 for Michael, including the revolutionary COBUILD project of the 1980s.

To many readers of *Lexikos*, Sue and Michael are of course well known. Back in September 1997, thus just two years after the establishment of Pan-SALB, and in order to prepare the future lexicographers of the 'soon-to-be established' eleven National Lexicography Units (a process finalized in March 2001), Sue Atkins and Michael Rundell, joined by Edmund Weiner, presented the two-week-long (monolingual) SALEX Lexicographical Training Course. This course was organized by Penny Silva, then Executive Director of the Dictionary Unit for South African English in Grahamstown, South Africa, currently Director of Editorial Projects for the *Oxford English Dictionary* in Oxford, United Kingdom. One year later, in September 1998, Sue and Michael returned to South Africa, for the first week of the follow-up (bilingual) AFRILEX-SALEX Lexicographical Training Course, organized by AFRILEX at the University of Pretoria.

While both Sue and Michael had already been involved in teaching courses for publishing houses for about a decade, the magnitude of the South African experience (with 35 participants at SALEX and 50 at AFRILEX-SALEX) surely contributed to their decision to turn their efforts into an annual event. Indeed, three years later, and joined by the computational linguist Adam Kilgarriff, they began their LEXICOM series, an intensive five-day workshop in lexicography and lexical computing. The detailed notes drawn up for the South

African courses, and perfected and complemented over the years for the LEXI-COM courses, eventually 'morphed into' (p. ix) OGPL.

The present commentator has been in the fortunate position to be a fellow traveller for a decade now, with attendances of the 1998 (Pretoria) and 2001 (Brighton) courses as a participant, and the 2005 (Brno), 2006 (Opatija), 2007 (Brno) and 2008 (Barcelona) courses as a guest presenter of the dictionary writing system *TshwaneLex*. One thing is certain: Sue and Michael kept the absolute best from their course notes, and added even better material on top.

OGPL, then, clearly has some of its roots in South Africa. As if to celebrate this fact in its own way, it is undoubtedly befitting that the University of Pretoria conferred an honorary doctorate on Sue Atkins on 5 September 2008 (see pp. 412-416 in this issue of *Lexikos*).

### 3. A gold and diamond-studded trophy

The main **strength** of OGPL is that it has a crystal-clear goal, and that the authors have managed to stay focused on that for over 550 pages. That goal is a detailed description of how today's monolingual and bilingual *corpus-driven commercial learners' dictionaries* are compiled. Although most reference works are actually adaptations of existing ones, Atkins and Rundell have decided to go all the way, and present every single step 'from the first twinkle in the eye of the publisher' (p. 18) through to the completion of the central dictionary text. Given the background of the authors, it will not surprise that the description of the compilation of a monolingual learners' dictionary is for an *English* one, and that that of a bilingual learners' dictionary is for an *English to French* one (and only that direction). Atkins and Rundell also decided to focus on the production of *paper* dictionaries for *human* users. Electronic dictionaries and NLP applications are mentioned where appropriate only.

The main **novelty** of OGPL is the minute description (which runs through and frames the entire book) of the so-called 'two- and threefold lexicographic process' — for monolingual and bilingual dictionaries respectively — used to compile corpus-driven dictionaries from scratch. Ironically, about this Atkins and Rundell state: 'we make no special claim for the methodology we outline in this book [...] What we describe here is what has worked well for us over a number of years' (p. 2). On the contrary, the whole point of the book is to make a strong case for exactly that approach.

Following an introductory chapter, the book is divided into **three parts**: the first deals with 'Pre-lexicography' (six chapters), the second with 'Analysing the data' (two chapters), and the third with 'Compiling the entry' (three chapters). This in itself is an extremely bold move. Recall that the aim was to describe the making of a dictionary from a corpus, yet the first part (i.e. before any analysis is being done) runs over an astonishing 250 pages. Rather than start with fireworks, and instead of discussing one topic at a time, the authors opted for a multilayered approach, whereby the same topic is revisited time

and again as the argument progresses. This results in a very large number of forward and backward cross-references. In §7.2.7.2, for example, one finds the following **forward cross-references** to sections in Chapters 9, 10 and 12: 'Choosing examples is a very important part of entry-writing, and is further discussed in §9.2.4, §10.8 and §12.3.3' (p. 225). Or, for the reverse, in §12.2.5 the following **backward cross-references** are found to sections in Chapters 6, 7 and 9: 'The vocabulary types represented by the linguistic labels are introduced in §6.4.1.4; the way the labels function is explained in §7.2.8; and their use in the database is set out in §9.2.9' (p. 496). There is thus a real need to read everything and it is hard to lift out sections, as one may miss a point (or two, three).

In **Chapter 1**, the 'Introduction', the book and its topic are briefly introduced. It can also be read as a 'thank you/acknowledgement' section. From the first sentence onwards, it is stressed that OGPL is about 'creating a dictionary' (p. 1), and by implication thus not about publishing, testing, analyzing, theorizing, etc. An important point is the following: 'You learn about lexicography by doing it, by training other people to do it [...], and by talking about it with colleagues' (p. 9). To any aspiring lexicographers: OGPL can thus best be used when put into practice.

**Part I**, 'Pre-lexicography', is the longest part with six chapters. The first of those is **Chapter 2** on 'Dictionary types and dictionary users'. Following an introductory section in which the entire process line of dictionary production is sketched, this chapter deals with the types of dictionary, the types of dictionary user, and then, combining the two, a section on tailoring the entry to the user who needs it. §2.4.1, on how to do the latter for monolingual dictionaries, is a nice, gentle analysis within the framework of this chapter's title. From the equivalent for bilingual dictionaries, one will do well to remember the following right from the start: 'the bilingual dictionary is more complex, and less amenable to clear explanations, than all but the most scholarly and sophisticated of the monolinguals' (p. 43).

The topic of **Chapter 3** is 'Lexicographic evidence', or reformulated: How to make sure that your dictionary is reliable? Here, *introspection* (with the required snipe at Noam Chomsky on p. 49) and *informant-testing* are contrasted with the use of *citations* and *reading programmes*. Although each of those is valuable, the type of evidence this book puts first and foremost is of course that found in large collections of electronic texts, commonly known as *corpora*. OGPL covers their design and the data collection in detail, and also sketches processing and annotation issues. The coverage is informative at all times, full of excellent advice, as in: 'never offer to pay for permission to include a text, and never agree to such a request from a copyright holder. Once money starts changing hands (even if for a single text in a single corpus), a precedent would be established that could have fatal consequences to corpus-creation efforts worldwide' (p. 83).

**Chapter 4**, titled 'Methods and resources', is a mixed bag. In the middle of this chapter the two chief 'resources' are discussed: a *corpus query system* (CQS) on the one hand, and *dictionary writing software* (DWS) on the other. The CQS

that is singled out is the Sketch Engine, with its linked lexical profiling tool, the Word Sketch. Regarding DWSs, Atkins and Rundell write: 'Two of the best and most widely used packages are IDM's Dictionary Production System [...] and the TshwaneLex dictionary compilation software' (p. 114). The last two sections of this chapter look into the *Style Guide* and the use of *template entries* (i.e. lexical-set compiling): 'As a rule, the members of a lexical set pose the same kinds of lexicographic problem and should be handled in the same way in a dictionary' (p. 124). The section on template entries is exceptionally well executed, and is one of the highlights of OGPL.

The opening section of this chapter focuses on the 'method' advocated by the authors. This method is revolutionary, as the compilation route is not straight from corpus to dictionary, but rather, a *twofold lexicographic process* whereby a database editor will first analyse the corpus and populate a pre-dictionary database. In a second step, a dictionary editor will then synthesize those findings in order to produce a dictionary. This is the process for monolingual dictionaries. For bilingual dictionaries, an extra step is needed in-between, hence the characterization as a *threefold lexicographic process*, whereby a translator populates the pre-dictionary database with (selected) translations. Without a good understanding of this process, OGPL cannot be used successfully: Part II (Chapters 8 and 9) is entirely devoted to the first step ('analysis'), while in Part III, Chapter 10 is devoted to the second step ('synthesis') for monolingual dictionaries, and Chapters 11 and 12 to the second and third steps ('transfer' and 'synthesis') for bilingual dictionaries.

In **Chapter 5**, 'Linguistic theory meets lexicography', the sense relationships *hyponymy*, *synonymy*, *meronymy* and *regular polysemy* are contrasted with *complementary antonymy*, *polar antonymy*, *directional antonymy* and *converseness*. The central section in this chapter, however, is the presentation of *Frame Semantics*, about which the authors say: 'the application of this theory to practical lexicography results in [an] approach to lexicographic relevance [...] which helps lexicographers to identify useful facts in corpus texts' (p. 144). In other words, Frame Semantics is the core linguistic theory which is advanced in OGPL to analyze corpus data. The last section of the chapter is on *lexicographic relevance*, which is actually an introduction to Chapter 9.

In a more traditional framework, **Chapter 6**, on 'Planning the dictionary', would be referred to with *macrostructure* decisions. The aims are: 'in §6.1 to clarify some of the basic terms and concepts; in §6.2 to name and describe the different kinds of words and phrases which you have to be able to recognize in corpus data and which can be dealt with as headwords in a dictionary; in §6.3 to set out the various large components of a print dictionary (the actual dictionary text and other material); in §6.4 to explain features you have to consider when deciding the words to include in the dictionary; in §6.5 to look at the other main decisions that have to be made about the headword list; and in §6.6 to describe the principal types of entry that are to be found in most current dictionaries' (p. 161). In the process, a large amount of (new) vocabulary is introduced and defined. The basic terms and concepts are *tokens*, *types*, *lemmas*, *lexi-*

*cal units (LUs)*, and *headwords* (§6.1). The types of lexical items discussed include: *lexical words* (nouns, adjectives, verbs, adverbs, interjections); *grammatical words* (prepositions, conjunctions, pronouns, auxiliary verbs, determiners); *alphabetisms, acronyms, contractions*; *bound affixes, productive affixes, combining forms*; *fixed and semi-fixed phrases* (transparent collocations, fixed phrases, similes, catch phrases, proverbs, quotations, greetings, phatic phrases); *phrasal idioms*; *figurative, semi-figurative and functional compounds*; *phrasal verbs*; and *support verb constructions* (§6.2). The constituent parts of a dictionary are of course: *front, mid- and back matter*, plus the *A–Z entries* (§6.3). The headword list proper consists of: *common words* (for which one needs to consider wordclass, lexical form, lexical structure, and vocabulary types (which leads to linguistic labelling)) and *proper names* (place names, personal names, and other names) (§6.4). In order to organize the headword list one needs to know more about: *alphabetization, syllabification, and homographs* (§6.5). Lastly, the types of entry are: *standard lexical, abbreviation, grammatical, and encyclopedic* (§6.6). The above enumeration serves as an example to show the dedication of the authors: every possible term and concept that is needed is also covered.

In a more traditional framework, **Chapter 7**, on 'Planning the entry', would be referred to with *microstructure* decisions. In analogy with Chapter 6, a very large number of terms and concepts are presented, all of them adequately illustrated by means of marked-up components of dictionary excerpts. In the section on labelling, Ogden and Richards's 'meaning triangle' is a delight. Tucked away in this chapter's back pocket, namely in subsection §7.2.11, is a basic introduction to electronic dictionaries.

In **Part II**, 'Analysing the data', Atkins and Rundell 'get down to business' (p. 261). They do this in two phases, first focusing on the identification of word senses of polysemous lemmas in Chapter 8, and then fleshing out those senses (or more correctly *lexical units*) in Chapter 9. As pointed out above, at this point no dictionary is being produced: a language (*in casu* English) is analyzed in detail, and the results of that (corpus) analysis are stored in a pre-dictionary database.

**Chapter 8**, titled 'Building the database (1): word senses' is a core chapter, which is no doubt why it was singled out to be offered as a free sample of OGPL on the publisher's website. The task can be described as follows: 'Meanings exist in infinite numbers of discrete communicative events, while the senses in a dictionary represent lexicographers' attempts to impose some order on this babel' (p. 311). Of the various 'theories' offered, one is especially relevant: *prototype theory* — which is complemented by discussions of *polysemy* and *homonymy, lexical semantics* and *motivated polysemy* (i.e. specialization, regular polysemy, figurative uses, metaphorical sets, conceptual metaphor, and metonymy), as well as a string of other useful ideas. Despite of all this, the authors point out: 'Learning about these ideas won't necessarily make the process of identifying word senses any easier, but you will tackle the job with greater confidence if you understand the underlying systems, and you will be better equipped to make good judgments in the more marginal, less clear-cut cases'

(p. 294). OGPL certainly offers a wide array of practical strategies for successful *word sense disambiguation* (WSD), strategies which will hopefully bring lumpers and splitters closer to one another — at least at the analysis (pre-dictionary) stage.

The bulk of **Chapter 9**, 'Building the database (2): the lexical unit', deals with *grammar in the pre-dictionary database* (§9.2.5), with a focus on (English) verbs, nouns, adjectives, and adverbs. For each of these wordclasses, the lexicographically relevant co-constituents are first listed in detail, then it is shown how those constructions can be retrieved from the corpus, and finally it is illustrated how examples may be selected to support each instance. Those four subsections are followed by what must be the absolute highlight of OGPL, namely a detailed ten-page-long 'case study on using complements in sense analysis, and in recording them' (p. 349). The headword being studied is the verb 'cook', and unlike the four 'theoretical' descriptions which preceded this subsection, the correct order is used here, viz. (1) analysing the corpus lines, (2) identifying the constructions to be recorded, and (3) recording complementation in the database. Note thus the order of (1) and (2). The beauty of this case study is that it is clearly shown how analyzing concordances in Frame Semantics 'Light' terms gives the valency patterns of each lexical unit of the headword. A pity indeed, then, that this gem is tucked away deep into this chapter (under §9.2.5.5) and that no cross-references whatsoever to it are to be found anywhere. It is not even listed under the entry 'frame semantics' in the book's Index. The rest of this chapter basically tells the reader how other database fields (such as headword, wordclass, meaning, examples, multiword expressions, collocates, patterns, labels, cross-references, and comments) are dealt with.

**Part III**, 'Compiling the entry', is 'the moment of truth' (p. 383). This section is perhaps best introduced with two quotations from Chapter 9: 'A database entry should be at least two or three times bigger than the final dictionary entry' (p. 322). 'Take the case of the word *coffee*. This word, in the sense of the substance, either solid or liquid, is uncountable; however when it is used to denote a single unit of coffee (in a cup, for example) it becomes countable; a similar phenomenon occurs when the word is used to mean 'type of coffee'. These three senses would normally be considered as different LUs in the database: what you do with them in a dictionary depends (as always) on the type of dictionary' (pp. 337-338). Dictionary editors will thus study all the facts that have been recorded by the database editors, and 'distil' from that exactly that type of dictionary which is deemed appropriate for the envisaged target users. A new team of dictionary editors may resynthesize the facts for a new edition in a slightly different way (which may also be accompanied by changes to the Style Guide), yet another team may extract a very different type of reference work altogether from the same database, whilst the largest ramifications will of course be seen when a second language is added to the database in order to produce a bilingual dictionary. Recall in this regard, for example, the various treatments of 'strike gold' in Section 1.

**Chapter 10**, 'Building the monolingual entry', concludes the discussion of the creation of a monolingual learners' dictionary. While most information during this stage may 'simply' be selected from the pre-dictionary database ('synthesis'), two new elements surface: presentation and definition writing. The first involves decisions on how to distribute the information, or thus what to do with, and where to place multiword expressions, run-ons, dictionary senses, grammar and labels. The second is a truly new task, which is why the bulk of this chapter (a massive fifty pages, or about 10% of OGPL) is concerned with how to write definitions that match the corpus-derived facts. Examples that had already been collected may at that point be selected from the database to support and complement the newly written definitions. All along, a close eye is also kept on the user profile and the Style Guide, while template entries are used whenever appropriate to ensure consistency.

**Chapter 11**, 'The translation stage', deals with stage two ('transfer'), of the threefold lexicographic process which characterizes the compilation of a bilingual dictionary. Here, translators 'work through each LU (both word senses and multiword expressions), adding target-language (TL) translations, going forward and back over the entry, and seeing which TL word seems to fit best as the first, or 'direct', translation — in essence, the word that suits most of the contexts before them. Then they decide which of the remaining contexts (those which the direct translation doesn't fit) are important enough to be kept in the entry, and translate the headword in these contexts' (p. 465). Important concepts that come into play here are *context-sensitive* vs. *context-free* translations (the latter is the aim of the dictionary), and *denotation* (literal meaning) vs. *connotation* (figurative meaning).

**Chapter 12**, 'Building the bilingual entry', mirrors Chapter 10 in more than one respect. Here, the bilingual dictionary editor will edit the entries from the translated pre-dictionary database, completing stage three ('synthesis') of the threefold bilingual process. The crucial difference with the monolingual effort — at least within the framework advocated by OGPL — is that 'dictionary senses' in a bilingual are not really senses of the headword at all, but simply the most user-friendly way to structure the material' (p. 500). As if to shatter one's beliefs even more, OGPL further claims that: 'It is naïve to think that you can lift stretches of corpus text, translate them, and produce really useful examples' (p. 507). After over five hundred pages of excited build-up, one ends up throwing out both dictionary senses and real corpus examples. All good things must come to an end.

#### 4. Technical aspects

If anything, Section 3 demonstrates the enormous ground covered by OGPL. Inevitably, such a mammoth undertaking also resulted in some infelicities. Those are looked into below.

#### 4.1 On the register and background

To those of us used to a strict academic **register**, the informal style of OGPL is immediately noticeable. Not only do very many sentences start with 'And', 'But', or 'So', large sections truly read like a novel, while the reader is often (but inconsistently) addressed. Atkins and Rundell use 'easy' language throughout, shunning any unnecessary metalanguage that is all too common among those who think that there is such a thing as 'theoretical lexicography'. Every now and then a 'hard word' is dropped, however, as the selection in Addendum 1 shows.

Written in English, the primary language of study for the monolingual effort is also English (although a single Spanish definition is flashed, p. 440). French is added to the equation for the bilingual effort (with one case of Finnish however, p. 41). OGPL, then, is fully embedded into the English and European cultural world, so mentions of say Westminster (p. 273), or the contributions of the Greek Aristotle (p. 437) and the Swede Linnaeus (p. 437), are all taken for granted. Specifically for lexicography, it is assumed that every reader already knows who Samuel Johnson (pp. 5, 48, etc.), James Murray (pp. 50, 52, 275), Noah Webster (p. 52), or A.S. Hornby (p. 455 (and p. 449)) are and what they did. Similarly for some lexicographic resources: the BNC is first mentioned on p. 41 (without any forward cross-reference to what it is on pp. 58-61); WordNet is only 'dropped' (p. 82). Readers also need some level of linguistic background to understand conventions such as '+HUMAN' (p. 155), the meaning of terms such as 'attributive', 'predicative', or (especially relevant in French) 'subjunctive' (all p. 221), or say 'Grice's maxim of quantity' (p. 456). All of these are not necessarily problematic; it depends on the audience (and the **background** of that audience) one wishes to reach.

#### 4.2 On the general structure

As is clear from Section 3 above, the overall structure, with the main chapters distributed across three parts, is very effective. Early on in each of the main chapters one also finds a **schematic overview** summarizing that chapter's contents and structure. This is an excellent idea (although the 'Preliminaries' are missing in Figg. 7.1 and 8.1 (> Fig. 6.1) while Fig. 10.1 is not referred to).

Unfortunately, in several places the **hierarchy** (and the related numbering system) went wrong. For instance, §2.1.2 should have been §2.1.1.2; §7.2.8.8 should not have been on the same level as the seven previous subsections; §8.3.4-6 should have been §8.3.3.1-3; §8.3.7 should have been §8.4; the current §8.4 then changes to §8.5 (with the current heading at §8.5 best dropped). Related, sections called 'Conclusions' are sometimes not conclusions at all, as is for example the case with §8.6 and all its subsections: this reads like an add-on, with far too much new information for it to be a conclusion. The same is true for the material on p. 431, which does not belong to the conclusion of §10.5.8.

Perhaps mimicking the latest range of learners' dictionaries, OGPL is printed in two **colours**: black (as well as some grey tints) and blue. This is especially effective when **dictionary sections** are shaded for ease of reference. The entire background colour of the so-called 'boxes' is also blue. **Boxes** are used to provide background information, and (in Chapter 9) for core tables and discussion points. Unfortunately, some of the boxes are not referred to (namely Boxes 3.1, 4.1, 7.1, 7.2, 7.3 and 9.7). While some of the extra information in such boxes is indeed truly excellent (e.g. Boxes 4.2 and 10.1), at times the material seems to be close to meaningless (e.g. Box 4.3). Also set in blue, is the right-pointing arrow (➔) which precedes all so-called 'rule of thumb practical tips' (p. 229).

Another excellent feature are the various lists of **pros and cons** to help readers make decisions for themselves, such as the enumeration of the advantages and disadvantages of citations (pp. 51-52), the pros and cons that deal with whether the 'first cut' in microstructure design should be based on grammar or meaning (pp. 248-249), or the options relating to multiword expressions as headwords (pp. 254-255). In most cases, however, the options are not made explicit and need to be 'derived' from the arguments presented, as in the insightful discussion of the use of full sentence definitions (pp. 441-443), or the presentation of the source and characteristics of good examples (pp. 455-461).

A true textbook feature comes in the form of **exercises**. While every chapter presents those at the end, they are missing from Chapters 1 (of course), 3, 4 (but the last section may be seen as one), and 5. The exercises are mostly simply straightforward applications of the points that were covered. Each chapter is formally concluded with a **Reading** section (for which, see Section 5.2 below).

### 4.3 On the multilayered build-up

As pointed out in Section 3 above, the presentation of the two- and threefold lexicographic process automatically led to a multilayered build-up of the entire line of argument. While this often works out well, there are some annoying consequences. Chief among those are cases of unnecessary **repetition**. The concept of a 'lexical set', for example, is first defined on p. 123, then again on p. 139, and yet again on p. 490. Large swathes of Chapter 12 are an especially bad example: pp. 484-488 are a near-calque of pp. 385-389 a hundred pages earlier. Likewise, §12.1.2 mimics §10.1.2, and §12.1.3 mimics §10.1.3. Surely, a simple cross-reference together with a summary of the differences would have sufficed (as was resorted to in the last paragraph of §12.2.1). Repetition sometimes also occurs within a single chapter, as for the two 'bullets' as well as the 'valency patterns' on p. 148 and again on p. 155. Some types of repetition result in a time warp, as in 'a day or two' (p. 508) becoming 'a week or two' (p. 513).

The **order** of the introduction of certain terms and lexicographic devices is not always correct either. For example, although the (undefined) concept 'lexical unit (LU)' had already been used a number of times, it suddenly takes central stage in the discussion of template entries (pp. 124 *fn*), but one only finds a

(first) short explanation in footnote 2 on p. 131, and a (second) full explanation on pp. 162-163. Of course, it is not always easy or even possible to introduce one aspect at the time, but the way the term 'canonical form' was handled on p. 168 — in a brief footnote that includes a cross-reference to the fuller explanation — is certainly more effective for a textbook. A similar reasoning can be made for the term 'lemma', which is used on p. 150, but only defined on p. 162, or for the term 'itemizer' on p. 323, defined on p. 371.

The same holds for some **symbols** and **metalinguage** in quoted dictionary articles. Two examples: The function of the (gender) letters '*m*', '*f*', '*fpl*', etc. should have been explained from their first appearance (in Fig. 6.14, p. 194) onwards. The (informality) markers '\*' and '\*\*' should also have been explained from their first appearance (in Fig. 7.6, p. 212) onwards, rather than (only partly) on p. 232, and (fully) on p. 498.

While some **building blocks** of the argument are indeed nicely stacked on top of one another (compare for instance Fig. 5.25 with Fig. 8.8), some blocks do not seem to fit. §9.3, following a sixty-page-long §9.2, is just 0.3 pages long, and amounts to one backward and two forward cross-references. Likewise, §12.2.3 seems to be just a place holder for §12.3.1. At other times, blocks, such as the 'dictionary pilot study' (p. 196), appear unexpectedly, once, and are never returned to.

## 5. References and cross-references

References and cross-references come in many different forms in OGPL. In the front matter, one finds a list of abbreviations which refer to dictionaries, and another one for an assortment of concepts; each chapter is concluded with a Reading section; the main text includes classical references to the Bibliography while the different building blocks cross-refer to one another; and in the back matter one finds a Bibliography plus an Index. References and cross-references hold everything together; unfortunately, they are sometimes a bit inconsistent. This is discussed below.

### 5.1 Abbreviations

On p. x, a list of **dictionary abbreviations** is presented. All of them are indeed used in OGPL, but, given there seems to be only one instance of CALD (on p. 389) one could as well have done without that dictionary abbreviation. The abbreviations are, however, not always used, and are sometimes written out in full: OALD on p. 22, MWC on p. 310, MW-3 on pp. 312, 417, 426, ODE on p. 400. The system used to refer to these dictionaries is not consistent either (e.g. '*CRFD-1998*' in Fig. 7.6 vs. '*CRFD-5 (1998)*' in Fig. 7.7). Quite surprisingly, not a single of these dictionaries has been provided with full bibliographic details. The same is true for the many other dictionaries, written out in full, that

are referred to in the text. Not one made it into the Bibliography.

On pp. xi and xii a list of '**other abbreviations**' is presented. One of them does not seem to be used anywhere in the text (VP), another one was already included with the dictionaries (OED), while quite a number are missing (L1, L2, N, XML). Several are in any case written out in full in the text (e.g. IPA on pp. 29 and 37, MLD on p. 410 n7, DWS on pp. 358, 489, 495 and 497).

## 5.2 Reading section

At the end of each chapter, a list of references is included to relevant background material. That Reading section is always divided into '**Recommended reading**' (R) and '**Further reading on related topics**' (F), sometimes complemented by a section '**Websites**' (W).<sup>1</sup> Quite surprisingly, for some Reading sections not a single source has been referred to from the preceding text (Chapters 7, 11, 12); for others just a handful (Chapters 9, 10-F). Two sources appear in both the R- and F-list, which is a contradiction (Rundell 2002 on p. 316, and Hanks 1987 on p. 464). One is missing from the Bibliography (Melčuk [sic] 2000 on p. 316).

## 5.3 Main text

If I have seen further it is by standing on ye shoulders of Giants. (Isaac Newton, Letter to Robert Hooke, 15 February 1676)

In order to be able to write such a revolutionary textbook as *OGPL*, Atkins and Rundell had to build on earlier work. At times, however, the text cries out for references. Proper and appropriate references to such works would not have diminished their own standing. Some examples: to the use of computers in the dictionary-making process, on pp. 3 and 112: **Urdang 1966, 1984**; to the published proceedings of the First Fulbright Colloquium, on p. 32: **Ilsion 1986**; to *CLAWS*, on p. 90; to an analysis of the 'synonyms' shut and close, on pp. 134-135: **Lewandowska-Tomaszczyk and Hanks 1996**; to semantic prosody, on p. 304 (although that term is not mentioned there, only on pp. 321, 375-376): **Sinclair 1987, 1998**; to the Theory of Norms and Exploitations, on pp. 309, 312, 397: **Hanks 2004**; or to 'the probable and the possible', on pp. 314, 398: **Hanks 2001**. The authors could have referred more often to their own work as well. To give one example: The gem described in Section 3 (Chapter 9) above, is based on **Atkins 2002**.

While it is understandable that **references** to novels (pp. 47, 73 n22, 270, 297, 408) have not been included in the Bibliography, and while it is true that one can simply google for *Google* (p. 48) or *Wiktionary* and *Wikipedia* (both p. 422), it would have been helpful if the source of all the concordances (or the name of the corpora used) had been provided throughout (where do, for instance, the examples on pp. 270, 279, 295, 297, etc. come from?). Or, again,

should one always simply resort to *Google* to find the source of a spoken transcript such as the one on p. 78, or to see who is being quoted in the 'Consignia anecdote' on p. 375? (The sources are CNN and *Industry News* respectively.) Saying that a quotation is from Johnson is also not good enough: those on pp. 5, 9 and 178 are from Johnson (1747), while the one on p. 448 is from Johnson (1755). Similarly for the quotation of Fillmore on p. 410; it is from Fillmore (2003).<sup>2</sup>

The different building blocks of the main text are held together by literally hundreds and hundreds of forward and backward **cross-references**. Most of these cross-references were placed in the running text itself, others have been relegated to footnotes for no obvious reason. Even so, some sections cry for more cross-references (so, at the bottom of p. 208 for instance, one would have liked to see a mention of §10.6.3.1).<sup>3</sup>

#### 5.4 Bibliography

The Bibliography does *not* bring together all the bibliographic details of the material referred to in OGPL. One reason was given in the paragraph on 'references' in the previous Section 5.3. Another reason is that references that are only mentioned in footnotes and not in the running text have been provided with full bibliographic details in those very footnotes (pp. 48 n3, 52 n5, 55 n10, 56 n12, 74 n24, 226 n13, 293 n20 (2 x), 369 n16, 401 n5, 419 n11, 422 n14, 436 n22).<sup>4</sup> Full bibliographic details may also, and only, be found in boxes (Box 3.1) and even in the running text (p. 400). It is not immediately clear why this user-unfriendly option was chosen.

#### 5.5 Index

A ten-page-long Index concludes OGPL. Regrettably, it is a subject index only; there is no author index. All the subjects and page numbers covered indeed appear in OGPL, but the coverage seems far from exhaustive. Firstly, numerous concepts have not been included at all. A random sample follows, including selected page numbers: active 506; faux amis 475; *pertainim* 305, 339, 343, 510; semantic prosody (304,) 321, 375-376; stratified sampling 64, 66, 68; passive 506.<sup>5</sup>

Secondly, often one or more page numbers are missing for those terms that are covered. A small sample: back matter +506; Brown Corpus +80; connotation +468-469; denotation +469; electronic dictionary +477; frame semantics +349 *fn* (esp. 354-355); IPA +29; priming +270, +297. The latter, 'priming', is an interesting case. On p. 270 and again on p. 297, the term is used without being defined, nor with any forward cross-references to where it is eventually defined (on p. 307). While it is acceptable not to give the meaning away too early, in order to 'prime' the readers, all instances *should* have been covered in the Index nonetheless.

The third type of problem with the Index is that some terms are treated in an idiosyncratic way. One example will suffice. While 'definiendum' (used on pp. 421, 440 and 435) is not included in the Index, 'definienda (form or function of) 439-440' is. It is not clear why the plural form was used as an entry for the Index, especially that it does not appear as a plural on pp. 439-440.

Lastly, the Index could have been put to good use in order to somewhat reduce the very large number of references and cross-references throughout the text, and especially in footnotes.

## 6. Errors and suggestions for improvement

Errors are inevitable in any project, so too in OGPL. Addendum 2 brings together around seventy (types of) errors. If the comprehensiveness of OGPL is taken into consideration, that number is not so dramatic, yet given all of them could and should have been picked up by the publisher, they remain regrettable. In Addendum 2, the errors were classified according to whether they are general or specific. In the second group, errors in quoted dictionary text, concordances, graphs and boxes are looked into. The first group presents the remaining typos, as well as problems with inflection and syntax, capitalization, and punctuation, and also lists inconsistencies, font and layout issues.

Addendum 3 contains some suggestions for improvements, again subdivided into subsections: more conventional formulations, likely better options, sections to improve upon, incorrect statements, and additional (recent) references.

## 7. Discussion points

There can be no question that for this type of textbook, the two **best** candidates joined forces. B.T. Sue Atkins, with her background of working as a lexicographer, manager and consultant for the great publishing houses of Collins/Robert and Oxford/Hachette, attended to the bilingual component, while Michael Rundell, with his experience of working as a manager for Longman and Macmillan, was responsible for the monolingual component. During their careers they certainly also came into contact with the best computational lexicographers. OGPL, then, was bound to be a hit. Their strength, however, also encapsulates some weaknesses.

Obviously, not 'everything' is (nor can it be) covered in OGPL. Very little if anything is said about **management** and time lines, for example, and tools to help with that.<sup>6</sup> Open any of the modern learners' dictionaries, on any page, and one wonders how long each of those articles took to be compiled. Or more down to earth: Fig. 9.10 summarizes just one LU of the verb *watch*, so one shudders at the thought of doing this type of analysis for every single sense of every single word of a language. On how to handle this, the authors could have

given us more information, as they have had to deal with it time and again.

Specifically for bilingual lexicography, the reader may also wonder how the **reverse side** should be compiled. Is one to start afresh, with a brand-new analysis of the second language, or could one hope to 'reverse' and thus reuse large sections of the side already compiled? Plus, in either case, how is one to cope with the 'reversibility principle'?<sup>7</sup> Again, some coverage of this would have been welcome.

Dictionaries of the type being described easily require over a hundred **contributors**, with many more collaborators in addition to 'database editor', 'translator' and 'dictionary editor', as any dictionary title page will confirm. How is one to organize this? What is best outsourced (function words, subject-specific terms, illustrations, grammatical sketches, all extra matter, etc.)? As the textbook stands, one may end up with a perfect central section, without any paper to wrap it in.

One drawback to the near-exhaustive presentation of the creation of a dictionary's central section is that the process intimidates. Surely, one knows all the secrets to become a lexicographer now, but also knows one cannot possibly assemble the materials. For the sake of the lesser fortunate, the options should (and could) have been downplayed every now and then. With regard to **corpora** and a CQS, why was the obvious not even mentioned, namely to simply dump plain text files into say *WordSmith*? Conversely, true corpus building for many (non-English) languages involves more than is covered by OGPL. For those without much or even any digital data, for example, scanning of printed data requires smart *optical character recognition* (OCR) 'tricks'. (While making a corpus from an oral language requires even more; but we would not explore this.) OGPL's near-exclusive focus on developments in and around the British Isles also failed to take account of some important work done elsewhere. Still on the issue of corpora, p. 79 boasts, with emphasis added: 'The New Corpus for Ireland — collected in 2003–4 — is an *early example* of a corpus which includes web data.' Hardly impressive, one would say, as in mid-October 1999, the present commentator had already built a corpus of Swahili sourced entirely from the Web (cf. e.g. De Schryver 2002: 268).

With regard to the selection of **examples** to populate the pre-dictionary database, OGPL states as ideal, with numbering added: '(1) It's short. (2) It provides an 'informative context' for the headword [...] (3) It has no words in it that are more difficult to understand than the headword. (4) It doesn't include words at variance with the register, style, region, etc. of the headword. (5) It doesn't contain the name of a real person, living or dead. [...] (6) It has no foreign words in it' (p. 330). If one compares this with the various excerpts of the pre-dictionary database in OGPL, one notices trespasses against more or less every point. Examples 1a and 2a in Fig. 9.13 for example are *very* long, while example 6b in the same Fig. 9.13 mentions Yeltsin. Compare this with (1) and (5). It seems as if OGPL here mixes the pre-dictionary database with the final dictionary.

Cases such as these may actually be a sign that the dictionary-writing process known as the **'two- and threefold lexicographic process'** — as *described*, and *with* the suggested software — has actually never been implemented. This is a strong statement, but one based on an impression that runs right through the reading of OGPL. If one goes back to the software that is being presented to support the dictionary-writing process (p. 103), one sees that two types are offered: a CQS and a DWS. Instantly, one not only wonders how data from the CQS ends up in the DWS (the 'dictionary database'), but more importantly, one needs to know where the 'pre-dictionary database' (which stores the analysis) fits in. The ensuing discussions of CQS (§4.3.1) and DWS (§4.3.2) do not give any hint. Each subsequent mention of 'database' in OGPL was carefully scrutinized (amongst others on pp. 123, 126, 157, 174, 264, 268, 317, 318, 319, 336), but to no avail. On p. 319 for example, one reads: 'Database and dictionary differ in structure as well as size', so one wonders which tools 'transform' that structure. On p. 336 a database extract is even shown (as text in a word-processor table), but one wonders why that extract is not rather a screenshot, as is customary when demonstrating software features. The one and only 'clue' is found on p. 358, where the 'pre-dictionary database' being populated seems to be equated with the 'dictionary writing system'. Each next mention of 'database' (such as on pp. 470 or 480 *fn*, both ideal candidates for screenshots) remains vague on this issue. Therefore, while one should not doubt that OGPL's advocated two- and threefold lexicographic process is indeed feasible, the way in which it is described, using just a CQS and a DWS, has not yet been implemented as such.<sup>8</sup> Once it has, it would be informative to see several case studies of words 'followed through' from corpus to (monolingual and bilingual) pre-dictionary database to final (monolingual and bilingual) dictionary articles.

## 8. A powerful achievement

The time is past when the lexicographer, using only her own judgement and knowledge of ... (Atkins ... 2008: 201)

... cricket would construe the expression in these terms (... and Rundell 2008: 297)

Despite the fact that there is some room for improvement, it should be clear from the analysis that this does not alter the fact that *The Oxford Guide to Practical Lexicography* is a most powerful lexicographic book-length achievement. Above all, it is witty throughout, as when Patrick Hanks is referred to as Johnson's 'twenty-first-century counterpart' (p. 274) or when the modifier that precedes Anna Wierzbicka is 'reliably contrarian' (p. 434). Rare are cases where science and art meet. Here is an instance!

## Endnotes

1. It is not clear whether or not the listings are meant to be in alphabetical order. Some are (Chapters 2, 4, 7, 8 (with one error), 9, 10 (with one error), 12), one is not (Chapter 1), while for others the R section is not, while the F section is (Chapters 3, 5, 6, 11). Also, the alphabetical order used in the Reading sections is generally different from that in the Bibliography for family names that start with a particle (i.e. 'de' and 'van'). For consistency, 'eds.' should not be used on p. 257.
2. For the sake of consistency, all references should mention the family name only, so '~~D.A.~~ Cruse (1986, 2004)' on p. 282, '~~Deirdre~~ Wilson (forthcoming)' on p. 285, '~~George~~ Lakoff and ~~Mark~~ Johnson's seminal *Metaphors We Live By* (1980)' on p. 290, '~~Charles~~ Fillmore (2003: 268)' on p. 409, and '~~Janet~~ Whitcut (1988: 53)' on p. 449. Also, on p. 44 one of the dates is wrong: 'Martin and Al ~~1988~~' → 1990; what is not yet published is better referred to with 'forthcoming', on p. 96: 'Prinsloo ~~2008~~' and 'Rundell and Atkins ~~2008~~'; while, on p. 464, it is not clear who the author is of pages '14-24' and '93-96' in 'Cowie (1987a)', given the latter is an edited collection.
3. One cross-reference, on pp. 210-211, is to a source outside the textbook.
4. The details for Biber 1990, on p. 67 n17, have been forgotten, however.
5. Rather than the dichotomy 'active/passive' (p. 506), 'encoding/decoding' seem to be the preferred terms elsewhere in OGPL. The latter pair is covered in the Index.
6. One of the measurement and management tools built into TshwaneLex in this regard is for example the Ruler tool (cf., for a theoretical background, De Schryver 2005).
7. Here too, TshwaneLex offers dedicated tools to help reverse either side of a bilingual dictionary, and to ensure that the reversibility principle is adhered to (cf. e.g. De Schryver 2008: 286 n6).
8. Similarly, OGPL is also vague on whether one is dealing with one or two Style Guide(s) during compilation (cf. e.g. pp. 359 vs. 363). It's probably two.

## References

### Dictionaries

- Atkins, B.T.S. and P. Robert (Eds.). 1978. *Collins-Robert French–English, English–French Dictionary / Robert-Collins dictionnaire français–anglais, anglais–français*. Glasgow: Collins / Paris: Le Robert.
- CRFD-8 (2006) = Atkins, B.T.S et al. 2006<sup>8</sup>. *Collins-Robert French Dictionary*. Glasgow: HarperCollins Publishers.
- MED-1 (2002) = Rundell, M. 2002. *Macmillan English Dictionary for Advanced Learners*. Oxford: Bloomsbury Publishing Plc.
- MED-2 (2007) = Rundell, M. 2007<sup>2</sup>. *Macmillan English Dictionary for Advanced Learners*. Oxford: Macmillan Education.

### Other sources

- Atkins, B.T.S. 2002. Then and Now: Competence and Performance in 35 Years of Lexicography. Braasch, A. and C. Povlsen (Eds.). 2002. *Proceedings of the Tenth EURALEX International Con-*

- gress, EURALEX 2002, Copenhagen, Denmark, August 13–17, 2002: 1-28. Copenhagen: Københavns Universitet, Center for Sprogteknologi.
- Atkins, B.T.S. and M. Rundell.** 2008. *The Oxford Guide to Practical Lexicography*. New York: Oxford University Press.
- BNC. 2008. *British National Corpus* [online]. <http://www.natcorp.ox.ac.uk/>.
- CLAWS. 2008. *CLAWS Part-of-Speech Tagger for English* [online]. <http://ucrel.lancs.ac.uk/claws/>.
- CNN. 30 July 2007. *Larry King Live — Triple-Murder Victims Mourned* [online]. <http://transcripts.cnn.com/TRANSCRIPTS/0707/30/lkl.01.html>.
- COBUILD. 2008. *Collins Birmingham University International Language Database* [online]. <http://www.cobuild.collins.co.uk/>.
- De Schryver, G.-M.** 2002. Web for/as Corpus: A Perspective for the African Languages. *Nordic Journal of African Studies* 11(2): 266-282.
- De Schryver, G.-M.** 2003. Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16(2): 143-199.
- De Schryver, G.-M.** 2005. Concurrent Over- and Under-Treatment in Dictionaries — The *Woordeboek van die Afrikaanse Taal* as a Case in Point. *International Journal of Lexicography* 18(1): 47-75.
- De Schryver, G.-M.** 2008. Why Does Africa Need Sinclair? *International Journal of Lexicography* 21(3): 267-291.
- Fillmore, C.J.** 2003. Double-Decker Definitions: The Role of Frames in Meaning Explanations. *Sign Language Studies* 3(3): 263-295.
- Google. 2008. Google Search Engine [online]. <http://google.com/>.
- Hanks, P.** 2001. The Probable and the Possible: Lexicography in the Age of the Internet. Lee, S. (Ed.). 2001. *ASIALEX 2001 Proceedings, Asian Bilingualism and the Dictionary*: 1-15. Seoul: Yonsei University, Center for Linguistic Informatics Development.
- Hanks, P.** 2004. The Syntagmatics of Metaphor and Idiom. *International Journal of Lexicography* 17(3): 245-274.
- Hanks, P.** 2008. Lexical Patterns: From Hornby to Hunston and Beyond. Bernal, E. and J. DeCesaris (Eds.). 2008. *Proceedings of the XIII EURALEX International Congress, Barcelona, 15–19 July 2008*: 89-129. Sèrie Activitats 20. Barcelona: Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada.
- Hoey, M. and M.B. O'Donnell.** 2008. Lexicography, Grammar, and Textual Position. *International Journal of Lexicography* 21(3): 293-309.
- Ilsou, R.F. (Ed.).** 1986. *Lexicography: An Emerging International Profession*. Fulbright Papers 1. Manchester: Manchester University Press.
- Industry News*. 17 June 2002. Dragon Brands Defends the 'Consignia' Re-Brand [online]. [http://www.brandintelligence.com/live/articles/dragon\\_defends\\_the\\_consignia\\_rebrand.asp](http://www.brandintelligence.com/live/articles/dragon_defends_the_consignia_rebrand.asp).
- Johnson, S.** 1747. *The Plan of an English Dictionary*. Edited by Jack Lynch [online]. <http://andromeda.rutgers.edu/~jlynch/Texts/plan.html>.
- Johnson, S.** 1755. *Johnson, Preface to the Dictionary*. Edited by Jack Lynch [online]. <http://andromeda.rutgers.edu/~jlynch/Texts/preface.html>.
- Kilgarriff, A.** 1996. *BNC Database and Word Frequency Lists* [online]. <http://www.kilgarriff.co.uk/bnc-readme.html>.
- Kilgarriff, A.** 1997. Putting Frequencies in the Dictionary. *International Journal of Lexicography* 10(2): 135-155.

- Lewandowska-Tomaszczyk, B. and P. Hanks.** 1996. Completive Particles and Verbs of Closing in English. Weigand, E. and F. Hundsnurscher (Eds.). 1996. *Lexical Structures and Language Use, Volume 2*: 89-103. Tübingen: Niemeyer.
- LEXICOM. 2008 *Lexicography MasterClass — Lexicom Workshops* [online]. <http://www.lexmasterclass.com/>.
- Sinclair, J.M.** 1987. The Nature of the Evidence. Sinclair, J.M. (Ed.). 1987. *Looking Up, An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary*: 150-159. London: Collins ELT.
- Sinclair, J.M.** 1998. The Lexical Item. Weigand, E. (Ed.). 1998. *Contrastive Lexical Semantics*: 1-24. *Current Issues in Linguistic Theory* 171. Amsterdam: John Benjamins.
- TshwaneLex*. 2008. TshwaneLex Dictionary Compilation and Terminology Management Suite [online]. <http://tshwanedje.com/tshwanelex/>.
- Urdang, L.** 1966. The Systems Designs and Devices Used to Process the Random House Dictionary of the English Language. *Computers and the Humanities* 1: 31-33.
- Urdang, L.** 1984. A Lexicographer's Adventures in Computing. *Dictionaries: Journal of the Dictionary Society of North America* 6: 150-165.
- Wikipedia*. 2008. Wikipedia — The Free Encyclopedia [online]. <http://www.wikipedia.org/>.
- Wiktionary*. 2008. Wiktionary — The Free Dictionary [online]. <http://www.wiktionary.org/>.
- WordSmith*. 2008. WordSmith Tools — Lexical Analysis Software for the PC [online]. <http://www.lexically.net/wordsmith/>.

**Addendum 1:** 'Hard words' in *The Oxford Guide to Practical Lexicography* (with Page = page number in OGPL; Freq., Word, POS and Files = frequency, word (i.e. type), part of speech, and number of files the word occurs in — all in the hundred-million-word BNC, as tagged by CLAWS (cf. Kilgarriff 1996), this to illustrate their infrequent use; this list obviously does not include words from citations, dictionary quotations or corpus lines)

Page	Freq.	Word	POS	Files	Page	Freq.	Word	POS	Files
43	325	amenable	aj0	245	179	2	siphoned	aj0-vvn	2
52	11	tetchily	av0	10		2	siphoned	vvd	2
68	120	creditable	aj0	102		2	siphoned	vvd-vvn	2
69, 434	37	egregious	aj0	32		20	siphoned	vvn	19
76	107	conscientiously	av0	96	285	77	amelioration	nn1	49
91	6	aggregating	aj0-vvg	4		1	amelioration	nn1-np0	1
	2	aggregating	nn1-vvg	2		1	ameliorations	nn2	1
	20	aggregating	vvg	18	285	—	pejoration		
95	47	unimpeachable	aj0	41	320	17	pre-ordained	aj0	17
157	—	anisomorphic				1	pre-ordained	aj0-nn1	1
166	60	chaff	nn1	44	323	57	multifarious	aj0	52
	2	chaff	nn1-np0	2	369	36	prescient	aj0	32
	12	chaff	nn1-vvb	10	426	75	epithet	nn1	64
168	8	phatic	aj0	7		3	epithet	nn1-vvb	3
168	69	nigh	av0	67	430	208	epistemological	aj0	54
	1	nigh	unc	1	437	2	ergative	aj0	2
	1	nigh/	unc	1	452	158	stricture	nn1	26
						1	stricture	nn1-vvb	1
						1	stricture(s)	nn2	1

**Addendum 2:** Errors in *The Oxford Guide to Practical Lexicography*

#### A. General

##### A.1 Typos

- 'a bilingual English–French dictionary' → a **unidirectional** bilingual ... (p. 24 n4)
- 'a bilingual English–French dictionary' → a **bidirectional** bilingual ... (p. 24 n5)
- **unshaded** → shaded (p. 209)
- Mel'čuk → Mel'čuk (p. 257)
- 'no hits at all **on** the BNC' → ... in the BNC (p. 412)
- 'So don't **to** try to force ...' → ... don't try to ... (p. 416)
- 'Meijs, W. ...' → Meijs, W. ... (p. 524)

##### A.2 Inflection and syntax

- '*adjudicate* sometimes **take** a direct object' → *adjudicate* sometimes takes ... (p. 60)
- 'Template entries are 'pro forma' entries for use by the lexicographers writing either database or dictionary.' → ... writing either **a** database or **a** dictionary (p. 123, plus similar cases)
- 'Apresjan (1973) described the semantic components which **gave** rise to ...' → ... give rise to ... (p. 139)

- 'rumours *circulate*, storms *rage*, heart *beats*' → ... a heart *beats* (p. 152)
- 'in this section we discuss the **factor** that influence these choices' → ... the factors that influence ... (p. 431)

### A.3 Capitalization

- **EURALEX** is written in caps, all other -LEXes are not (p. 10)
- 'be used as **Linguistic data**' → ... linguistics data (p. 51)
- Word Sketch(es) ↔ **w**ord **s**ketch(es) (eg. pp. 107 *fn* vs. 369-370)
- 'those who treat dictionary users as if they are **Lexicographers**' → ... as if they are lexicographers (p. 436)
- **A**bbreviation → abbreviation (p. 531)

### A.4 Punctuation

- '(senses 1, 3, 4, and **5**).**This** technique ...' → ... and 5). This ... (p. 38)
- 'lexical-profiling software, for example (discussed in §4.3.1.5) **only**' → ... **only** (p. 61)
- **etc** vs. etc. (e.g. pp. 104 vs. 105)
- '... and Moon 1998, **Mel'čuk** 1998 presents ...' → ... Moon 1998. Mel'čuk 1998 ...' (p. 167)
- ). → .) (pp. 182, 451)
- .) → ). (p. 280)
- .) → .) (p. 461)
- 'including a Kiswahili–English one: (<http://africanlanguages.com/swahili/>)' → ... Kiswahili–English one (<http://africanlanguages.com/swahili/>) (p. 257)
- 'two broad semantic groupings: 'relationships' and documents' (or ...' → **documents**' (p. 302)
- 'Part-specific PP-specific' → last ' should not be in italics (p. 331)
- 'given in Box 9.#, **together** ...' → ... in Box 9.#, together ... (pp. 331, 343, 347)

### A.5 Inconsistencies, fonts and layout

- 'wording of definitions' vs. 'wording of the definitions' (e.g. pp. 35 vs. 38)
- bullets that do not match the ensuing text (e.g. pp. 57 *fn*)
- inclusion vs. exclusion of 'www' in URLs (e.g. pp. 83 vs. 96)
- fonts of figures need to be streamlined (e.g. Fig. 4.6, p. 108 vs. Fig. 4.10, p. 125)
- space needed between first and second group in Fig. 5.10 (p. 137)
- text needs to be broken up in extra paragraphs (pp. 215, 303, 373, 417, 461)
- 'collocate type' [sic] and 'collocate-type' is mentioned twice, in two different bullets (p. 322)
- meaning line should be dropped in Fig. 9.2 (p. 326)
- **I** *just stood* ...' → 'I should be in italics (p. 335)
- 'one reason **for** **Mr Major's reluctance** to ...' → 'for' should be in bold (p. 341)
- construction 9 from Fig. 9.13 not in Box 9.3 (pp. 341-342 vs. 339-340)
- indentations for most quoted material are missing in Chapter 10 (pp. 393 *fn*)
- not used when pages face each other, so may be dropped: '(cont.)' (p. 432)

## B. Specific

### B.1 In quoted dictionary text

- missing space after headword (Fig. 6.14)
- missing spaces after tilde (Figg. 2.3, 7.6, 7.14)

- no space required after tilde (Fig. 12.7)
- missing space after label (Fig. 7.37)
- missing space after plus sign (Fig. 7.29)
- missing space after colon (Fig. 12.3)
- missing dot after POS (Fig. 7.16)
- missing dot in ellipsis (Fig. 7.28)
- bits should not be bold (Figg. 2.3, 6.2, 6.17, 6.19, 12.6)
- bit should be in bold (Fig. 12.10)
- wrong font (Fig. 12.10, 12.20)
- wrong colours (Figg. 7.3, 12.5, 12.22)
- missing colour (Fig. 12.7)
- misplaced colour (Fig. 7.17, 12.21)
- dot should be comma (Fig. 12.5)
- typo in 'nous retrouverons une solution' → nous trouverons ... (Figg. 2.7 and 2.8)
- missing space in 'ask sth(for sth)' → ask sth (for sth) (Fig. 5.24)
- missing words (Figg. 7.17, 7.28)
- star marks should be on the same level (Fig. 7.6)
- OALD examples from Fig. 7.28 do not come from the fifth edition
- reference to the dictionary itself inconsistent (e.g. Figg. 5.6 vs. 6.14)

**B.2 In concordances**

- .. → . (Fig. 6.5)
- missing dots (Fig. 6.7)
- in general: the use of full stops is not consistent (e.g. p. 301)

**B.3 In graphs**

- the graph does not correspond with the text (and truth) for the stretch 1960s–1970s (p. 58)

**B.4 In boxes**

- use of '+' vs. 'with' inconsistent (Boxes 9.2, 9.3, 9.4)
- use of 'infinitive verb' vs. just 'verb' inconsistent (Boxes 9.2, 9.3, 9.4)
- alphabetical sorting not fully correct (Boxes 9.3, 9.4)

### **Addendum 3:** Suggestions for improvement in *The Oxford Guide to Practical Lexicography*

**A. More conventional formulations**

- 'the languages of southern Africa' → the Bantu languages (p. 2)
- 'superior number' → superscript number (p. 192)
- 'dedicated side columns' → Extra Column (p. 221)

**B. Likely better**

- Addition to the Reading section of Chapter 6: Kilgarriff 1997 (p. 199)
- It would have been a good idea to show concordances for 'happier' and 'happiest' as well in Fig. 9.15 (p. 345)

- In §9.2.10 one could have added that cross-references are also used between synonyms ('=') and antonyms ('><'), and to show words are related to one another ('<' and '>') (p. 378)
- 'The nouns shown in square brackets' → ... in italics (p. 470)

**C. May be improved upon**

- The reader may wonder: What happened to the lemmatization phase? (p. 90)
- The concepts 'homography', 'homophony' and 'homonymy' are not treated satisfactorily. (cf. esp. pp. 191-193, 280-282)
- Excessive prudence is shown in the treatment of 'offensive terms': in contrast to all other terms that need labelling, not a single example is shown. (p. 229)
- With reference to electronic dictionaries, OGPL claims: 'The field is developing so quickly that what we write today is almost guaranteed to be out of date by the time this book appears' (p. 238). This is a huge overstatement, as a comparison with even a five-year-old study of the field (De Schryver 2003) will easily show.
- Box 9.1: Where is the field for frequency? Also note that nothing whatsoever is said about the manipulation of frequencies in the database, even though statements about rare vs. frequent occurrences abound. (p. 321)
- With reference to a defining vocabulary (DV), one important issue is not mentioned, namely that expressed in number of senses, one *is* in fact dealing with most of the lexicon. (p. 449)
- While §§11.2.1 and 11.2.2 are arguably brilliant, §11.3 disappoints in its simplicity. (p. 468 *fn*)

**D. Incorrect statements**

- With reference to Fig. 7.28, one reads: 'The note at the head of the *OHFD into* entry gives the English-speaking user some general advice about how to put this preposition in French' (pp. 234-235). This is incorrect. What the 'local usage note' does is to give guidance on where in the dictionary users need to search for the translation equivalents of the various constructions with *into*.
- 'The homograph number is manually entered' → surely the software (a DWS?) automates this?! (p. 325)
- In the 'Introduction to Part III' it is said that a methodology was described in Chapters 8 and 9 to build a database (correct) and to translate that database (incorrect). The latter is actually the topic of Chapter 11 only. (p. 383)

**E. Additional (recent) references (for a second edition)**

- To complement the discussion of discourse structure and rhetorical features (p. 81): Hoey and O'Donnell 2008
- To complement the discussion of the coding systems lexicographers use(d) to indicate grammatical behaviour (p. 401, as well as Box 9.2 on pp. 332-333): Hanks 2008