An Analysis of
Practical Lexicography: A Reader
(Ed. Fontenelle 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: Intended as a companion volume to *The Oxford Guide to Practical Lexicography* (Atkins and Rundell 2008), Fontenelle's book aims to bring together the most relevant papers in practical lexicography. This review article presents a critical analysis of the success thereof, both in quantitative and qualitative terms.

Keywords: PRACTICAL LEXICOGRAPHY, DICTIONARY, MONOLINGUAL, BILINGUAL, QUANTITATIVE EVALUATION, QUALITATIVE EVALUATION, CITATION PATTERNS, CORPUS TOOLS, MEANING, DEFINITIONS, EXAMPLES, EQUIVALENCE, DICTIONARY USE, ENGLISH, FRENCH

Samenvatting: **Een analyse van *Practical Lexicography: A Reader* (Fontenelle 2008).** Bedoeld als achtergrondlectuur bij *The Oxford Guide to Practical Lexicography* (Atkins en Rundell 2008), is het doel van Fontenelle's boek om de meest relevante papers m.b.t. praktische lexicografie samen te brengen. Dit recensieartikel is een kritische analyse van het succes daarvan, zowel in kwantitatieve als kwalitatieve termen.

Sleutelwoorden: PRAKTISCHE LEXICOGRAFIE, WOORDENBOEK, EENTALIG, TWEE-TALIG, KWANTITATIEVE EVALUATIE, KWALITATIEVE EVALUATIE, CITATIEPATRONEN, CORPUS SOFTWARE, BETEKENIS, DEFINITIES, VOORBEELDEN, EQUIVALENTIE, WOOR-DENBOEKGEBRUIK, ENGELS, FRANS

1.   The Editor and the Reader

Think of Microsoft and computational lexicography, and a couple of names spring to mind. Ken Church is one of them, Thierry Fontenelle the other. The first will return below, the second is the editor of the book under review. In addition to his work as a Senior Program Manager with Microsoft's Natural Language Group (where he was responsible for the French lexical database used in a variety of natural language applications), Fontenelle is well known as an Associate Editor of the *International Journal of Lexicography* (IJL), as a Past President of Euralex, for the projects he led at the European Commission Translation Service, for his innovative research into collocations and semantic

networks, and for his contribution to bilingual dictionaries published by CUP. Fontenelle is currently the Head of the General Affairs Department at the Translation Centre for the Bodies of the European Union.

Invited by Sue Atkins to compile a companion volume to *The Oxford Guide to Practical Lexicography* (Atkins and Rundell 2008), henceforth OGPL, he proceeded with such speed that his 'accompanying' collection, *Practical Lexicography: A Reader* (Fontenelle 2008), henceforth PLR, was completed and published before OGPL. We opted to analyse these two works in their logical order, however, with a review of OGPL appearing in last year's *Lexikos* (De Schryver 2008), and a review of PLR herewith.

PLR is a collection of 23 papers. All of these papers have been published elsewhere, with the exception of the Introduction by Fontenelle, which provides an insightful overview of the whole topic, as well as a brief summary of each paper. Some of the papers are not easily accessible elsewhere and, in the opinion of this reviewer, well over half of them are required reading for any student with an interest in words, meanings, and dictionaries. The references for each paper have all been brought together into one single section at the end of the book,[1] and a combined author/subject index concludes PLR.

The core facts about each of the 22 papers as well as the Introduction are enumerated in the Addendum, to which reference is made for the details of various claims below. As may be seen from the Addendum, there are, altogether, 25 authors for the 23 contributions: Atkins and Kilgarriff have each authored or co-authored three; while Fontenelle, Grefenstette, Hanks, and Rundell are each author or co-author of two. The number of pages varies widely, from 4 pages (for Bolinger) to 39 pages (for Atkins and Varantola); the average being 14.8 pages. Following the Introduction, the 22 papers have been grouped into twelve parts, with either one, two or three papers per group.

2.   Reader statistics

This is not the first reader in lexicography/lexicology. Two of the better known ones are Hartmann's (2003) *Lexicography: Critical Concepts* (reviewed in De Schryver 2005), and Hanks's (2008) *Lexicology: Critical Concepts* (reviewed in De Schryver 2008a). Both are actually multi-volume anthologies (respectively three and six volumes), in contrast to the single-volume PLR. Their aim is also different: While both anthologies intend to provide an overview of the whole field (of lexicography/lexicology), the current reader sets out to support a specific textbook (OGPL). Cross-comparing various statistics for these three collections will nonetheless prove fruitful.

A first aspect of interest is to look at the dates of publication for each of the 23 texts in PLR. This is done in Figure 1, from which one sees that Fontenelle especially selected texts produced during the first half of the 1990s, and to a lesser extent from the following decade. In comparison, in Hartmann's lexicography collection, "an increasing number of influential texts were seemingly

written during the 1980s" (De Schryver 2005: 94). For Fontenelle, then, the lexicographic contributions that best support OGPL are written about a decade later. Compared to Hanks's anthology, which shows that lexicology has "continued to pick up momentum ever since [the 1950s]" (De Schryver 2008a: 421), an inverse trend may be noticed in the papers from the past two decades in PLR.
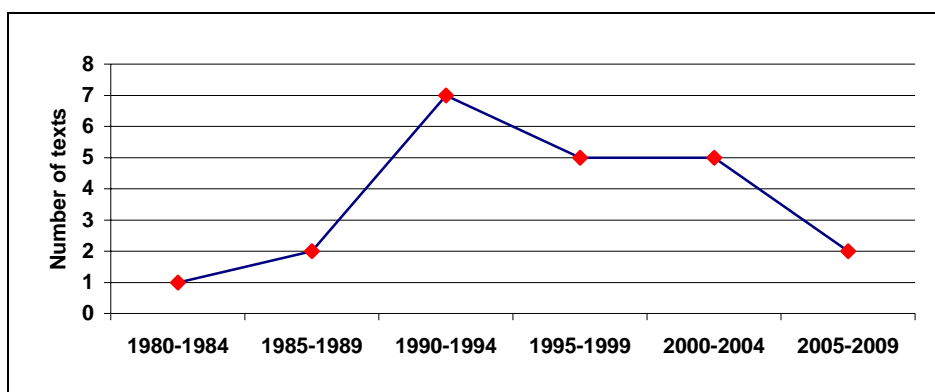
Figure 1: Distribution across time of the texts in *Practical Lexicography: A Reader* [not showing one mid-eighteenth-century text]

A second aspect of interest is a study of the sources of each of the texts in PLR. The distribution is shown in Figure 2, from which one may see that nearly 40% of the texts were initially published in journals (four, or nearly half, of them in IJL), and about 35% in conference proceedings (with as many as six, or three quarters, in Euralex proceedings). If anything, the large amount of material from IJL and the Euralex proceedings immediately delineates the field of study as Anglo-Saxon. In this it indeed supports OGPL, which "is fully embedded into the English and European cultural world" (De Schryver 2008: 431).
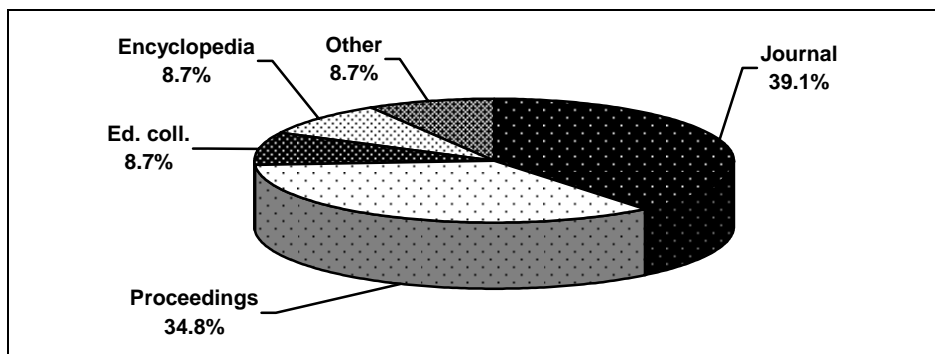
Figure 2: Sources of the texts in *Practical Lexicography: A Reader*

Comparing the distribution of the texts in PLR with the distributions of the texts in the two anthologies, one obtains the data listed in Table 1.

Table 1: Source distribution in three lexicography/lexicology collections

|	PLR (Fontenelle 2008)		Lexicology: CC (Hanks 2008)		Lexicography: CC (Hartmann 2003)	
	N	%	N	%	N	%
Journal	9	39.1	41	41.0	12	17.1
Proceedings	8	34.8	15	15.0	17	24.3
Edited collection	2	8.7	26	26.0	12	17.1
Dictionary / Encyclopedia	2	8.7	3	3.0	1	1.4
Book	–	–	13	13.0	18	25.7
Textbook	–	–	–	–	4	5.7
PhD	–	–	–	–	1	1.4
Other	2	8.7	2	2.0	5	7.1
	23	100.0	100	100.0	70	100.0

Clearly PLR's distribution is closer to Hanks's than to Hartmann's, with the book category entirely absent (where it was still 26% in Hartmann, and half of that, 13%, in Hanks). With an increased focus on journals and conference proceedings, rather than on books (and textbooks), this points to a lively and healthy research environment in practical lexicography.

A third aspect of interest is the answer to the question: "At what age does one write material worthy of inclusion in a reader of practical lexicography?" That age seems to be 49, compared to 47 in the lexicology anthology, and 51 in the lexicography anthology. The PLR average thus falls right in-between those of general lexicology and general lexicography — in line with what one would have expected, given the relatively high percentage of computational lexicographers (who tend to be younger than their colleague-lexicographers) in PLR.

Lastly, one can also compare the three collections with reference to the specific texts selected, as well as their contributing authors. Here the differences in overlap are striking. While only one text (Johnson 1747) and four authors (Cowie, Hanks, Johnson, and Varantola) are shared between Hartmann's collection and PLR, two texts (Church and Hanks 1989, and Kilgarriff et al. 2004) and as many as thirteen authors (Apresjan, Atkins, Bolinger, Church, Fellbaum, Fillmore, Fontenelle, Hanks, Kilgarriff, Miller G.A., Rychlý, Smrž, and Tugwell) are shared between Hanks's collection and PLR. With more than half of the PLR authors also represented in the lexicology collection, one is tempted to accord more value to the selections in these two, which depict a more unified field, rather than to the more esoteric selection found in the lexicography collection, which shows "a strong bias towards especially Asian authors and colleagues working in Asia" (De Schryver 2005: 94).

3.   Quantitative evaluation

This section presents a quantitative evaluation of PLR. The evaluation proceeds in two steps. Firstly, one can assume that the main purpose of a reader is not to present esoteric texts that the editor may happen to know; rather, the true core texts of a field ought to be brought together. From this it follows that quite a number of these texts should have attracted a substantial number of citations over the years. Bibliometrics, especially in lexicography, is still in its infancy (cf. De Schryver 2009a), but one rather relevant and freely accessible source in this regard is Google Scholar. With it, the number of citations for each of the 22 texts in PLR may be checked — the results of which are shown in Figure 3.
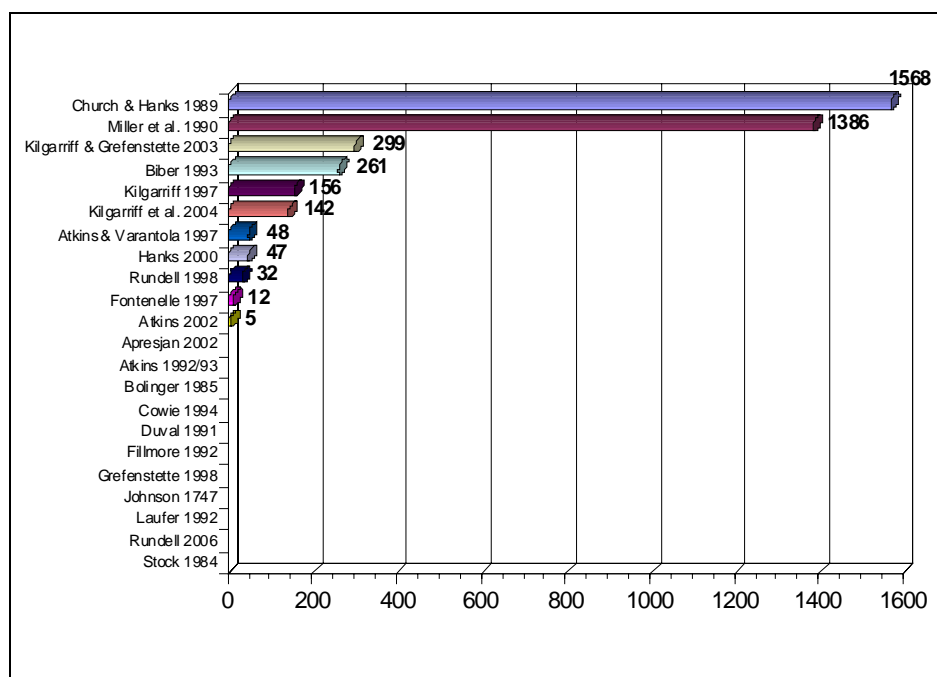
Figure 3: Number of citations for the texts in *Practical Lexicography: A Reader* (according to Google Scholar statistics on 5 August 2009)

As may be derived from Figure 3 (and thus according to the Google-crawlable material only!) half of PLR's contributions have indeed been cited (11 out of 22). The other half, however, has not, or at least much less (knowing that Google does not (yet) see 'everything').[2] Furthermore, the top six most-frequently cited are unsurprisingly contributions in the sub-field of computational lexicography — 'unsurprisingly', as the NLP community is (a) much larger than the lexicographic community, (b) prominently present on the Web, and (c) characterized by a much greater speed in communicating (and citing) research results. This

leaves us with just five truly (traditional) lexicographic papers that have been cited: Atkins and Varantola (1997), Hanks (2000), Rundell (1998), Fontenelle (1997), and Atkins (2002). If one merely looks at the author names, these papers from PLR simply *must* support the arguments in OGPL; although, of course, one cannot help but notice some circularity in the undertaking.

For the second step in the quantitative evaluation, one can take one step back to OGPL. In OGPL, each of the twelve chapters is concluded with a reading section, divided into 'Recommended reading' (R), 'Further reading on related topics' (F), and 'Websites' (W). One could assume that each recommended source would at least have been mentioned once in the chapter to which it belongs, but this is not the case, as may be seen from the statistics presented in Table 2.

Table 2: Referred vs. non-referred reading material in OGPL (per chapter, 'Ch.')

Ch.	Recommended				Further reading				Websites			
	Referred		Not ref.		Referred		Not ref.		Referred		Not ref.	
	N	%	N	%	N	%	N	%	N	%	N	%
1	2	40.0	3	60.0	1	25.0	3	75.0	5	71.4	2	28.6
2	1	20.0	4	80.0	5	23.8	16	76.2	0	0.0	1	100.0
3	6	75.0	2	25.0	10	55.6	8	44.4	1	11.1	8	88.9
4	2	66.7	1	33.3	1	14.3	6	85.7	0	0.0	6	100.0
5	6	66.7	3	33.3	12	32.4	25	67.6	1	25.0	3	75.0
6	2	50.0	2	50.0	2	6.9	27	93.1	0	0.0	1	100.0
7	0	0.0	5	100.0	0	0.0	51	100.0	0	0.0	2	100.0
8	6	50.0	6	50.0	17	35.4	31	64.6	–	–	–	–
9	0	0.0	5	100.0	7	10.0	63	90.0	–	–	–	–
10	7	87.5	1	12.5	17	23.6	55	76.4	–	–	–	–
11	0	0.0	4	100.0	0	0.0	19	100.0	1	50.0	1	50.0
12	0	0.0	4	100.0	0	0.0	37	100.0	–	–	–	–
	32	44.4	40	55.6	72	17.4	341	82.6	8	25.0	24	75.0

Of the five sources recommended at the end of Chapter 2, for example, only one was actually referred to in Chapter 2. Overall, there are 72 recommended references (including doubles), 44.4% of which have been referred to in the respective chapters. The number of sources referred to from the further-reading section is even lower: 17.4% overall (for a total of 413 references); with the website section falling in-between: 25.0% referred (for a total of 32 references).

While this mismatch was obviously a design feature by the authors of OGPL, one in which they departed from the academic practice of providing all relevant references throughout the text (cf. De Schryver 2008: 434-435), this leaves the present reviewer with the task to find an alternative way to check whether the texts in PLR are a good representation of the material referred to in OGPL. The only way to do so remains to check *all* references in OGPL, throughout the *entire* text. While reading through OGPL, a citation database was kept, recording not only the data summarized in Table 2, but also the

occurrence of each and every citation (C). Table 3 lists those references that were mentioned at least five times overall.

Table 3: Top section of the citation database built for OGPL, contrasted with the texts and authors in PLR

#	Author(s)	Year	R	F	C	Sum	Text in PLR?	Author(s) in PLR?
1	*Landau*	*2001*	4	2	6	12	–	–
2	Atkins	1992/93	3	2	1	6	✓	✓
3	Atkins, Rundell & Sato	2003	2	2	2	6	–	✓ ✓ –
4	Apresjan	1973	2	1	3	6	–	✓
5	*Cruse*	*2004*	2		15	17	–	–
6	Fillmore & Atkins	2000	1	3	1	5	–	✓ ✓
7	*Cruse*	*1986*	1	2	9	12	–	–
8	Geeraerts	1990	1	2	4	7	–	–
9	Bolinger	1965	1	1	4	6	–	✓
10	Hanks	1987	1	1	4	6	–	✓
11	*Aitchison*	*2003*	1	1	3	5	–	–
12	Johnson	1755	1		12	13	–	✓
13	Johnson	1747	1		7	8	✓	✓
14	Sinclair	2003	1		6	7	–	–
15	Kilgarriff, Rundell & Dhonnchadla	2007	1		5	6	–	✓ ✓ –
16	Moon	1987a	1		5	6	–	–
17	Stock	1984	1		5	6	✓	✓
18	Atkins, Clear & Ostler	1992	1		4	5	–	✓ – –
19	Bogaards	1990		4	1	5	–	–
20	*Hoey*	*2005*		3	7	10	–	–
21	Cowie	1999		3	2	5	–	✓
22	*Lakoff & Johnson*	*1980*		2	3	5	–	– –
23	Rundell	1998		2	3	5	✓	✓
24	*Taylor*	*1995*		2	3	5	–	–
25	Moon	1987b		1	4	5	–	–
26	Fontenelle	1996			5	5	–	✓
27	Hanks	2004			5	5	–	✓

When relative weight (R > F > C) is combined with number of citations, the most important source referred to in OGPL is (the second edition of) Landau's well-known textbook, with 12 references overall (4 x R, 2 x F, and 6 x C). This is followed by Atkins's (1992/93) DSNA paper (3 x R, 2 x F and 1 x C). Of course, Fontenelle could not have been expected to include entire textbooks or general linguistic monographs in his reader,[3] which immediately excludes numbers 1, 5, 7, 11, 20, 22 and 24 (all displayed in italics) from Table 3. Of the 20 remaining sources, four have been included in PLR (these are highlighted for ease of reference). While just four is not a very high number, the author representation is better: 11 are in (Atkins, Apresjan, Bolinger, Cowie, Fillmore, Fontenelle, Hanks, Johnson, Kilgarriff, Rundell, and Stock), versus 14 that are out (Aitchi-

son, Bogaards, Clear, Cruse, Dhonnchadla, Geeraerts, Hoey, Lakoff, Landau, Moon, Ostler, Sato, Sinclair, and Taylor).

One may now also proceed to analyse the reverse, namely, how does Fontenelle's selection compare to the citations in OGPL? The result of this comparison is shown in Table 4.

Table 4: Texts in PLR, contrasted with the citations in OGPL

#	Author(s)	Year	R	F	C	Sum
1	Atkins	1992/93	3	2	1	6
2	Johnson	1747	1		7	8
3	Stock	1984	1		5	6
4	Rundell	1998		2	3	5
5	Apresjan	2002	2		1	3
6	Cowie	1994	1	2	1	4
7	Atkins & Varantola	1997	1	2		3
8	Biber	1993	1	2		3
9	Fillmore	1992	1	2		3
10	Kilgarriff	1997	1	1		2
11	Hanks	2000	1		3	4
12	Kilgarriff & Grefenstette	2003	1		2	3
13	Bolinger	1985	1		1	2
14	Grefenstette	1998	1		1	2
15	Kilgarriff, Rychlý, Smrž & Tugwell	2004	1		1	2
16	Atkins	2002	1			1
17	Duval	1991		2		2
18	Rundell	2006		1	3	4
19	Laufer	1992		1	1	2
20	Miller GA, Beckwith, Fellbaum, Gross & Miller KJ	1990		1	1	2
21	Church & Hanks	1989		1		1
22	Fontenelle	1997		1		1

The following observations can be made. Firstly, each of the texts from PLR is mentioned at least once in OGPL. Secondly, all texts from PLR have either been recommended (R) and/or suggested for further reading (F) in OGPL. Thirdly, three texts from PLR are only listed once (Atkins 2002, Church and Hanks 1989, and Fontenelle 1997), which is meagre indeed. (After all, within the roughly fifty other items from the recommended lists, there are many with far more mentions in OGPL.) Including Atkins (2002) was not a good idea: A full rework and update of her paper was included as part of OGPL itself (as §9.2.5.5 in there). In the light of the citation patterns seen in Figure 3, it is truly astonishing that the paper by Church and Hanks (1989) only gets one mention in OGPL, so it is good news that it was 'rescued' for PLR. When Table 3 is compared with 4, it is also clear that item 26 in Table 3 would have been a better choice than Fontenelle (1997), but then it is the prerogative of the editor of a reader to select what they want to be remembered for.

All in all, then, the various aspects of the quantitative evaluation presented in this section indicate that about half the selection is warranted, both in terms of citation patterns and from the point of view of its aim to be a companion volume to the textbook it is to support.

4.   Qualitative evaluation

With the quantitative evaluation behind us, we are now in a position to present a qualitative evaluation. While each of the 23 contributions will be given attention, no attempt was made at balancing the length of each evaluation. This is not because of the varying length of the papers themselves (some of the shorter ones actually contain the most insightful thoughts), but simply because not all papers warrant the same attention from the point of view of being included in a reader on practical lexicography.

4.1   Fontenelle's *Introduction* (2008)

In his Introduction, Fontenelle both explains the genesis of his book and provides succinct summaries of each of the 22 texts he selected, successfully contextualizing each of them in the process, and providing additional references for further reading. Re-summarizing this part will only serve to perpetuate summaries of summaries, which cannot be the idea. Rather, two observations will be made. Firstly, Fontenelle claims that the structure of PLR reflects the structure of OGPL (p. 2).[4] While it is true that both OGPL and PLR contain the same number of chapters, twelve, that is where the similarity in structure ends, as there is hardly any correlation between the two series of twelve. What Fontenelle probably meant to say is that the *topics* selected for PLR parallel some of the *topics* developed further in OGPL.

Secondly, Fontenelle claims that, with reference to excellent papers in lexicography, "it must be acknowledged that many of these papers are often published in hard-to-access conference proceedings", and points out that his collection "attempts to meet the need for a coherent and easily accessible compilation of papers" (p. 2). It is not immediately clear whether or not the Euralex proceedings are really so hard to find, but what is verifiable is to check what can already be found in digital form. From the details listed in the Addendum, one sees that as many as 12 of the 22 texts are currently available online: 6 on the Internet, 1 through Google Books, and another 5 by subscription. Given that the Euralex proceedings are soon "to be made available free of charge through the Euralex website under a Creative Common[s] Licence agreement" (Bogaards 2009: 354), another 5 papers will easily become accessible at that point, bringing the total to 17 out of 22. As was pointed out in reviewing Hanks's lexicology collection (De Schryver 2008: 429-430), the future of readers and anthologies is online.[5]

4.2   Johnson's *Plan of a Dictionary of the English Language* (1747)

Fontenelle's collection, quite rightly, starts with Samuel Johnson's Plan. It is the first of three texts in Part I, *Metalexicography, Macrostructure, Microstructure, and the Contribution of Linguistic Theory*. Even though Johnson's Plan was written over 260 years ago, Johnson remains the only major literary figure in any language ever to have undertaken the burdensome task of compiling a dictionary — which he did alongside a prodigious outpouring of journalism, reviews, literary criticism, short biographies, and poetry. Hanks (2005: 265) remarks that, in the Preface to his dictionary (1755), "Johnson crisply addresses theoretical issues which were subsequently neglected for some two hundred years." Fontenelle has selected Johnson's less well-known "Plan of a Dictionary of the English Language", written eight years before the dictionary was published. A comparison of the two documents reveals some fascinating differences. In the Plan, Johnson announces that his chief intent is "to preserve the purity and ascertain the meaning of our English idiom" (p. 20). By the time the dictionary was published, Johnson had realized that natural languages cannot be fixed, but are governed by irresistible forces of change — an insight that to this day eludes certain conservative members of the academies of France, Italy, and Spain, as well as many other people who dislike linguistic change.

According to Johnson, a dictionary should provide guidance on at least orthography, pronunciation (he gives examples such as *tear* rhyming with *dare* vs. *tear* rhyming with *peer*), inflections, and etymology (a science which was in its infancy in Johnson's day), as well as the meaning of words. He has interesting comments on analogy ("speech was not formed by an analogy sent from heaven. It did not descend to us in a state of uniformity and perfection, but was produced by necessity and enlarged by accident" (p. 24)), syntax, phraseology, and several other matters. The Plan announces an intention that the dictionary should 'instruct the learner' as well as 'delight the critic' (p. 20). Both these objectives were to be amply fulfilled. The following quotation provides helpful guidance to metalexicographers, even a quarter of a millennium after it was written: "The unlearned much oftener consult their dictionaries, for the meaning of words, than for their structures or formations" (p. 20).

Johnson's Plan recognizes the problem of coverage of technical terms, but is rather ambivalent about how many should be included. In the event, the dictionary was to include many technical terms from 18th-century science, sometimes with a quotation from an authoritative source in place of a definition. Collectively, these terms provide an amazing insight into how much science has changed in 250 years: 18th-century scientific terms and explanations of terms are very often incomprehensible to modern readers without special training.

Concerning the order of senses of polysemous words, Johnson was well aware of the tendency of words to develop new senses through analogical or

metaphorical change. He offers (p. 26) the first known account of what came to be known as 'historical principles'. For example for *arrive*:

(1)   to reach the shore in a voyage (the "natural and primitive signification" [because *riva* in Latin means 'shore']).

(2)   to reach any place whether by land or sea (the "consequential meaning").

(3)   to obtain any thing desired, as in: he *arrived* at a peerage (a "metaphorical sense").

Furthermore, Johnson anticipates Sinclairian notions of phraseology and semantic prosody: "the word *arrive* […] cannot be properly applied but to words signifying something desirable; thus we say, a man *arrived* at happiness, but cannot say, without a mixture of irony, he *arrived* at misery" (p. 26).[6]

4.3   Atkins's *Theoretical Lexicography and its Relation to Dictionary-Making* (1992/93)

Much of what is said in this article is covered (often in a revised and improved form) in OGPL, so including it here may seem like unnecessary duplication of effort. Nevertheless, it does make a valuable contribution, especially for readers who only read PLR and not OGPL.

Atkins contrasts the single-volume "trade dictionary, a product created to be sold in the marketplace" (p. 31), with scholarly historical dictionaries. According to Atkins (p. 33), there are two steps to dictionary-making: analysis and synthesis. They may even be carried out by different groups of people. Analysis consists of "trying to discover as many relevant linguistic facts [about each word] as possible" (p. 33), by means of studying a large corpus and/or a collection of citations and/or pre-existing dictionaries and/or the lexicographer's own intuitions. Fortunately, the days in which lexicographers had to rely on their intuitions for evidence of linguistic facts about a word are mostly behind us, although for some languages (especially the undocumented ones), evidence based on the lexicographer's own intuitions still plays a large part. Gradually, large corpora of texts are being built for all languages, although in some cases — such as American English (see e.g. Hanks 2009) — dictionary publishers have not yet realized the need for corpus evidence. It should be borne in mind that publishers do not really care what the dictionary entries say, or whether what is said is true or false, as long as deadlines are met and the product meets its sales targets.

In the synthesis stage, the lexicographers write up the entries. Dictionaries are sometimes written by large teams, rather than by a single author, so a style guide, drawn up by senior editors, is needed to ensure consistency of treatment.

Atkins might have added a third stage: refinement. The wording of definitions or translations, as well as the choice of examples, is checked in dic-

tionaries far more often than in most other kinds of literary product, and changes — sometimes quite radical changes — are introduced at quite a late stage. The success of a dictionary in the market place may be quite radically affected by changes at the refinement stage, for the initial rough-hewn synthesis may not be clear or convincing to a general reader.

Atkins says (p. 34): "One clear, albeit indirect, contribution that the theoretical linguist may make to the synthesis process is to give the would-be lexicographer language skills and language awareness." However, as Atkins herself goes on to show, the criticisms levelled by academic linguists at dictionaries are often irrelevant and in other cases take no account of practical constraints such as space or the needs of users. Twenty years on, it is now clear that only some linguists can give the would-be lexicographer language skills and language awareness. Others merely confuse the issue and miss the point. Linguists such as Apresjan, Fillmore, and Halliday certainly have useful things to say to lexicographers, but others do not.

In a personal communicaton, Patrick Hanks mentions that, during the heyday of Chomskyan syntactocentrism (in the 1960s, 70s, and 80s), he regarded a degree in linguistics as an active disadvantage for a would-be lexicographer. Graduates with degrees in botany, classics, or literature made better lexicographers, for they had less to unlearn and more skill in discrimination. We may conclude from such attitudes by a well-known project leader in lexicography that even today would-be lexicographers need guidance on what sort of linguistics might improve their relevant language skills. On p. 45 Atkins seems to acknowledge this, for, after alleging that "lexicography will improve if more lexicographers read theoretical papers", she goes on to say: "the most helpful and common-sense papers on defining that I know are those written by practicing and practical lexicographers." PLR provides precisely the sort of selection of articles that Atkins in 1992/93 was pleading for.

On pp. 43-48 Atkins distinguishes 'internal facts' (facts about a word's spelling, inflections, and meaning) from 'external facts': "the editor's greatest problems with external facts lie in SYNTAGMATIC relationships" (p. 47). The truth of this important comment is now widely recognized and the problems that Atkins alludes to are beginning to be addressed in corpus-based syntagmatic research.

4.4   Apresjan's *Principles of Systematic Lexicography* (2002)

The Russian tradition in linguistics (Ščerba, Apresjan, Mel'čuk, amongst others) has always enjoyed an integral relationship between the study of lexis, pragmatics, semantics, syntax, and other aspects of language. It has never been bedevilled by a division between 'commercial dictionaries' and 'academic research'. Instead, dictionaries and research have gone hand in hand. A prime example is Apresjan, who is both a leading academician and practical dictionary maker. He was working without the benefit of corpus evidence, so he

sometimes got the details wrong, but his five general principles, outlined in this contribution, repay study by every would-be lexicographer. In brief, they may be paraphrased as follows.[7]

According to the first principle each language 'forces' its speakers to express specific meanings, such as in certain Russian sentences where one is forced to specify the manner of locomotion through verbs like *walking*, *flying* or *crawling*, as the use of the (more) general *leaving* results in doubtful constructions. The second principle insists on a perfect 'coordination' of dictionaries and grammars. The current discrepancy is convincingly exemplified with a discussion of the labelling of numerals as either 'nouns' or 'adjectives' in English dictionaries vs. their characterization as 'numerals' in their own right in grammars. In the third principle it is advocated that 'lexical classes' ought to be treated in full and described uniformly in dictionaries. Under the heading 'lexicographic types', factive and putative predicates are looked into as an example, leading to highly interesting oppositions (in English) such as 'knowledge has a source, but not a reason' vs. 'opinions have a reason, but never a source'. The converse verbs *buy*, *sell*, *pay* and *cost* are discussed as an instance of a 'lexical-semantic paradigm' under a second heading. Here, Apresjan anticipates Fillmore's theory of Frame Semantics. Moving from the macro- to the microcosm, the fourth principle stresses the importance of an exhaustive linguistic description of lexemes, while the fifth points to the need to pay attention to meaning interaction across language units.

Apresjan's text is a highly entertaining one, and the five principles of systematic lexicography that are defined and illustrated are indeed of great importance to practical dictionary making.

4.5   Biber's *Representativeness in Corpus Design* (1993)

Biber's article is the first of two that make up Part II, *On Corpus Design*. This text has attracted a considerable number of citations (cf. Figure 3), mainly from the computational linguistics community, it has been referred to from OGPL (cf. Table 4), and it is useful for corpus builders, *but* it is only indirectly relevant to lexicography, and in fact irrelevant to practical lexicography. While lexicographers of course need to know that the corpus evidence they are using contains a reasonably wide variety of different texts and text types, they do not need the sort of details on sampling procedures and corpus design that are to be found in this article.

The conclusion, too, is somewhat of a downturn: "the parameters of a fully representative corpus cannot be determined at the outset. Rather, corpus work proceeds in a cyclical fashion" (p. 86). Given the ease with which corpora are built these days, and especially the ease with which any number of sub-corpora may be combined and recombined at any point in time, Biber's text offers less than is intimated.

4.6   Kilgarriff and Grefenstette's *Introduction to the Special Issue on the Web as Corpus* (2003)

Kilgarriff and Grefenstette start their article with the important question "Is the Web a corpus?" Their answer is "Yes." Our answer is "Not really." The Web is a vast collection of texts, constantly growing and changing. A corpus is also a collection of texts, but it is a collection with a purpose. Insofar as that purpose is linguistic analysis, a corpus must be stable, for corpus analysis can only be effective if it compares like with like and is able to measure linguistic facts statistically. Thus, the Web is not really a corpus, the reason being that it is unstable. It is a resource from which many corpora can be built (i.e. the Web 'for' corpus building rather than the Web 'as' a corpus, a distinction introduced in De Schryver 2002) — just as it is a resource for all sorts of other purposes. For example, by using the Web, armchair linguists have the best chances of finding authentic evidence to support their pet theories, especially if they are too bizarre to be reflected in a small sample and therefore are not part of conventional usage. In contrast to this, the advantage of a real corpus — a stable corpus — is that it can be analysed statistically, enabling lexicographers and linguists to distinguish normal usage from abnormal (even if authentic) usage, or as Hanks (2008a: 228) put it: "authenticity alone is not enough: evidence of conventionality is also needed."

Kilgarriff and Grefenstette then present a very idiosyncratic account of the history of corpora, point out that the BNC is not big enough (no corpus linguist would disagree; but then, no one is arguing a case for reducing the size of corpora), with the subsequent sections of little relevance to practical lexicography, and in places rather tedious. So, this article (after all, merely an introduction to a special issue of the journal *Computational Linguistics*) is really disappointing, at least as far as lexicography is concerned. Here is why. Lexicography is concerned, among other things, with discovering and representing the conventional, normal use and meaning of each word in a language. The Web as corpus is a source of masses of evidence for word use of all kinds, but it is not a reliable source of evidence for conventional word use or significant collocations. It is therefore surprising that this paper has been included in a collection claiming to provide readings in practical lexicography.

In the opinion of the present reviewer, then, neither of the papers in Part II belong in this reader. They are not really relevant to lexicography, and they do not give a good overview of the *relevance* of corpus data to lexicography. In that sense, one could have hoped that the next contribution (cf. Section 4.7) would have obviated this problem, but unfortunately, it does not.

4.7   Fillmore's *'Corpus Linguistics' or 'Computer-aided Armchair Linguistics'* (1992)

Part III, *On Lexicographical Evidence*, consists of just one paper, Fillmore's. The paper opens with an amusing caricature of the mutual incomprehension be-

tween theoretical (generative) linguists and corpus linguists, and concludes by acknowledging that he "refuses to give up his old ways [involving introspection including the invention of evidence] but [nevertheless] finds profit in being a consumer of some of the resources that corpus linguists have created" (p. 105). The paper includes a full rehash of the well-known investigation by Fillmore and Atkins (1992) of the semantic frame for *risk*. This work was a precursor of the FrameNet database, in which the frame elements for all words participating in a given frame are identified. In the case of *risk,* the main frame elements are:

(1)   the person taking the risk — the **Protagonist**.
(2)   the **Harm** that the Protagonist might suffer.
(3)   a **Valued Possession** which the Protagonist puts at risk.
(4)   an **Act** performed by the Protagonist.
(5)   the **Goal** of the Protagonist.

Syntactically, (1) is the subject of the active verb *risk*. (2), (3), and (4) compete for the direct object slot, so only one of them can be explicitly present in any given sentence. (5), if present at all, is governed by the preposition *for*. Very often, the Goal is implicit rather than expressed. Nevertheless, semantically, all five arguments are implicitly present.

There is also (pp. 113-121) an extended and very interesting discussion of the word *home*. Both these case studies (*risk* and *home*) are highly thought-provoking. They are full of insights which contribute greatly to the understanding of a wide range of lexical items, and they may result (we hope) in due course in significant improvements in monolingual lexicography. They should be required reading for every novice lexicographer.

Two aspects of Fillmore's work may be criticized. One has already been mentioned, namely his view that corpus evidence is supplementary to evidence invented by introspection: He is not a supporter of *systematic* corpus analysis. The second is that he fails to distinguish evidence from interpretation: "Should it ever come about that linguistics can be carried out without the intervention and suffering of a native-speaker analyst, I will probably lose interest in the enterprise" (p. 122). Here, Fillmore seems to imply that corpus linguists want to abolish the use of intuitions for purposes of linguistic analysis. It is true that some extremists in the corpus-linguistics world believe that the right thing to do is to let language learners loose on concordances with WordSmith (Scott 2009) or some similar toolkit, and leave them to draw their own conclusions. An answer to such people is implicit in the paper by Laufer in this collection (cf. Section 4.15).

4.8   Hanks's *Do Word Meanings Exist?* (2000)

Hanks's article is the first of three texts in Part IV, *On Word Senses and Polysemy*. In it, Hanks is concerned with the question "Do Word Meanings Exist?" His answer is "Yes, but …" (p 133):

— Yes, but traditional dictionaries give a false impression. What dictionaries contain are (more or less inaccurate) statements of *meaning potentials*, not meanings.

— Yes, but only *in context*.

— Yes, but the meaning potential of a word consists of a cluster of *semantic components*, only *some of which are activated* when the word is used in context.

The article contains a corpus-based contrastive analysis of the two English nouns *bank*, in terms of their semantic components and semantic types, which Hanks associates with Pustejovsky's notion of the 'lexical conceptual paradigm'. It shows very clearly the systematic variability that characterizes word use in natural language such as English, and (by implication) the need for a new kind of lexicography that will link the prototypical phraseology of each word with a prototypical meaning.

4.9   Kilgarriff's *I Don't Believe in Word Senses* (1997)

Although the title of Kilgarriff's article may seem like a response to Hanks's question, this is misleading, for (a) it was published before Hanks's study, and (b) it is actually a statement (with which Kilgarriff agrees) by Sue Atkins dating from October 1994. While Hanks looked at meanings from the point of view of a seasoned writer and editor of monolingual dictionary definitions, Kilgarriff takes the NLP view, and is interested in automated word sense disambiguation (WSD). More particularly, Kilgarriff wants to find out whether the word sense division as seen in dictionaries or thesauri is of any use in NLP applications. His answer, unsurprisingly, is in the negative: "The set of senses defined by a dictionary may or may not match the set that is relevant for an NLP application" (p. 151). This outcome led to much further research, most of it undertaken under the umbrella of the Senseval project (see the Senseval website for more information, and Kilgarriff's Home Page for the numerous WSD publications).

   Nonetheless, as was the case with the two contributions in Part II, although relevant to the NLP community, it is not immediately clear how practicing lexicographers may benefit from this study, apart from knowing that the results of their labour cannot be put to good use for a purpose for which they were not intended.

4.10   Stock's *Polysemy* (1984)

Although rarely quoted (cf. Figure 3), Stock's paper is indeed a must-read for all practicing lexicographers. Published in 1984 and thus well before the COBUILD project heralded a revolution in lexicography — in the form of a very new type of learner's dictionary (Sinclair and Hanks 1987), and the first

book-length guide to making a dictionary from a corpus (Sinclair 1987) — this is the very earliest paper to draw attention to the importance of collocational analysis in actual corpus-driven lexicography.

Stock starts her paper by pointing out her dissatisfaction with Ayto's working method for sense division: "Distinct superordinate or genus words suggest distinct senses" (p. 153). While working on the COBUILD project, she noticed that it is "possible to disambiguate meanings from written material with minimal, and purely linguistic, context" (p. 156). With reference to an even earlier work by Jones and Sinclair (1974), she continues with: "It is clear that, in a large number of cases when working from concordanced citational material, an examination, sometimes even a fairly cursory examination, of the syntactic and collocational patterns in the environment of the node word (the word under analysis), clarifies which meaning is being used" (p. 156).

With ample corpus evidence at hand, she then proceeds to offer her own instructive method to determine isolable meanings of polysemous words. That method consists of three procedures: An analysis of the syntactic behaviour, a study of the collocational patterns, and a final check of each possible reading.

4.11   Cowie's *Phraseology* (1994)

Cowie's text is the first of two that belong to Part V, *On Collocations, Idioms, and Dictionaries*. Cowie is well known as the co-author of the two-volume *Oxford Dictionary of Current Idiomatic English* (Cowie and Mackin 1975; Cowie, Mackin and McCaig 1983), henceforth ODCIE, which is basically a dictionary of idiomatic phrases, based on an OED-style collection of citations, accumulated at vast and admirable expenditure of time and effort, rather than a corpus. A new edition of this classic work in lexicography, thoroughly revised in the light of corpus evidence, is overdue.[8]

The problem with ODCIE highlights the problem with the selected article: Cowie refers to an "accumulation of descriptive studies throughout the 1970s and 1980s" (p. 163), but these (meritorious as they may be) are based on introspection and citations, not on a corpus-based search for regularities. Much corpus-based work has been done in phraseology since this short article was written and it now appears rather dated.

4.12   Fontenelle's *Using a Bilingual Dictionary to Create Semantic Networks* (1997)

Fontenelle used the 1st edition of the Collins–Robert French–English bilingual dictionary to create a semantic network. He augmented it with lexical functions based on the lexicographic work of Igor Mel'čuk, more in particular the Meaning-Text Theory (MTT). The result is a bilingual lexical-semantic database that can be used and/or integrated with WSD programs, translation systems or corpus query tools. From an end user's point of view: "One of the ultimate

goals of lexical knowledge acquisition is to make it possible for a user to navigate within a lexical knowledge base through concepts and lexical relations" (p. 170).

Although this article is very representative of Fontenelle's own (earlier) research interests and work, one notices that (a) the resulting database is not publicly available (which calls into question its true research and exploitation potential), and (b) the relevance to practicing lexicographers can only be said to be indirect. Fontenelle himself suggests that his work "should be seen as a contribution to the study of lexical-semantic relations and, more specifically, of collocational knowledge" (p. 188).

4.13   Bolinger's *Defining the Indefinable* (1985)

If a prize were to be awarded to the highest quality/page ratio in PLR, it would definitely go to Bolinger's four-page squib, the first of two texts in Part VI, *On Definitions*. The following quote from the opening page (p. 193) should suffice to send everyone straight to Bolinger's text:

> Lexicography is an unnatural occupation. It consists in tearing words from their mother context and setting them in rows — carrots and onions and beetroot and salsify next to one another — with roots shorn like those of celery to make them fit side by side, in an order determined not by nature but by some obscure Phoenician sailors who traded with Greeks in the long ago. Half of the lexicographer's labor is spent repairing this damage to an infinitude of natural connections that every word in any language contracts with every other word, in a complex neural web knit densely at the center but ever more diffusely as it spreads outward. A bit of context, a synonym, a grammatical category, an etymology for remembrance's sake, and a cross-reference or two — these are the additives that accomplish the repair. But the fact that it is a repair always shows, and explains why no two dictionaries agree in their patchwork, unless they copy each other.

Brilliantly written, it is full of insight. In order to illustrate how much lexicographers destroy when they define, Bolinger then uses the suffix *-less* as his case study.

4.14   Rundell's *More Than One Way to Skin a Cat: Why Full-Sentence Definitions Have Not Been Universally Adopted* (2006)

Rundell's paper evaluates COBUILD-style full-sentence definitions; his point is that full-sentence definitions are appropriate only in certain circumstances, for example, the phrasal verb *to be laid up* in COBUILD-3 (Sinclair and Fox 1995):

> If someone **is laid up** with an illness, the illness makes it necessary for them to stay in bed.

COBUILD definitions are always longer than their conventional equivalents. Rundell's charge that sometimes they are often also unnecessarily verbose is

well justified. Rundell quotes the noun *retreat* from COBUILD-1 (Sinclair and Hanks 1987):

> A **retreat** is a change in your position when you have decided that you do not want to do what you have agreed or promised to do, usually because it has become too difficult, too expensive, or too embarrassing.

Rundell says with some justice: "longer definitions mean a heavier reading load (for readers whose linguistic resources are limited), and generally entail increased complexity. Thus the abandonment of traditional conciseness can bring new problems for users, who may go from the frying pan of unpacking a dense, formulaic definition to the fire of processing something two or three times longer" (p. 201). It was unfair of Rundell, however, to quote the 1st edition, as in COBUILD-3, the definition for this particular noun sense was dropped completely. The nearest thing to it is shown as a derivative of a verb sense, thus:

> **2** … When an army **retreats**, it moves away from enemy forces in order to avoid fighting them. *The French, suddenly outnumbered, were forced to retreat. | Retreating soldiers were dousing homes and shops with petrol and setting them on fire.* […]
>   Also a noun. *In June 1942, the British 8th Army was in full retreat.*

Furthermore, for balance, Hanks's seminal 'Definitions and Explanations' (1987), in which the rationale for the COBUILD defining style is explained, should have been included in PLR. As it stands, the massive fifty-page section on definitions in OGPL (pp. 405-452 in there) is far more successful in presenting pros and cons.

4.15   Laufer's *Corpus-based versus Lexicographer Examples in Comprehension and Production of New Words* (1992)

Laufer's text is the only one in Part VII, *On Examples*. In it, Laufer attacks one of the well-established practices of COBUILD lexicography, viz. the use of authentic examples only in the dictionary. She conducted research showing the following results:

(1)   Learners perform much better in acquiring understanding and learning to use new words if they are given a definition *and* examples, than if they are given examples alone (as in some corpus-based classroom work) — even if the examples are carefully preselected and sorted.

(2)   Examples invented by lexicographers tend to be more useful to learners than authentic examples taken from a corpus.

She argues that possible slight loss of naturalness is a small price to pay for improved comprehensibility. She has a point, and it is well supported by her research.

We might add, however, that there is a danger of confusing academic research with practical tool creation. A dictionary is a practical tool for learners, who do not want to be bothered with the niceties of academic disputes. On the other hand, Sinclair (1984) showed that the practice of invention of examples can seriously distort the patterns of conventional usage associated with every word. This latter point may be more important in an academic research context than in a language learning context, but of course as more and more data are made available and corpora grow larger and larger, it becomes easier and easier to obtain the best of both worlds, by selecting examples from corpus data that combine brevity and clarity with authenticity.

As with Rundell's contribution on definitions, which required a voice from the other side, Fox's (1987) 'The Case for Examples', in which she sheds light on the COBUILD approach to examples, should have been included in PLR for balance.

4.16   Rundell's *Recent Trends in English Pedagogical Lexicography* (1998)

Rundell's article is the only one in Part VIII, *On Grammar and Usage in Dictionaries*. As background reading to OGPL, it fulfils its role admirably well. In short, OGPL deals with the monolingual learner's dictionary (MLD), which is (a) a commercial product, and (b) for human consumption. OGPL is thus a textbook dealing with the production of real dictionaries, for use in real situations, by users with real needs. Those needs are both of the decoding as well as the encoding type. Providing dictionary users with encoding skills (an aspect formerly only found in teaching material of the non-dictionary type) is the first truly 'hard part' of modern MLD compilation. The use of corpus evidence since the mid-1980s has further led to the increased realization that the mere description of words in isolation is simply not sufficient: In order to convey meanings one needs to know more about the typical company words keep and the contexts in which they are used. Hence the greater attention to, among others, lexical collocations, multi-word expressions, syntactic environments and usage labelling. Getting all of this right is the second 'hard part' of modern MLD compilation. In evaluating PLR as a companion volume to OGPL, one thus wants to see contributions that inform these issues further. This selection by Rundell does so.

Rundell's text starts with the observation that "the pace of change has been rapid — driven by a combination of theoretically-informed innovation, astonishing technological advances, and the creativity of dictionary-publishers in response to the known and perceived needs of users" (p. 221). He then proceeds with an analysis of how much MLDs have changed since A.S. Hornby, focusing on their descriptive and presentational improvements. Although the overview ends with the state-of-the-art of a decade ago (naturally, as it was published in 1998), all that is said (including the predictions made) remain valid.

4.17   Atkins's *Then and Now: Competence and Performance in 35 Years of Lexicography* (2002)

Atkins's paper is the first of two in Part IX, *On Bilingual Lexicography*. It is a good paper, but even though longer than its embedded and reworked version in OGPL itself (pp. 349-359 in there), it does not have the admirable quality of the OGPL version (cf. De Schryver 2008: 429). As such, including it in PLR may indeed provide the extra bit of data and insight, but it is mainly of interest to forensic text criticism. The 'Then' part (missing from OGPL) is moreover very short (less than two pages). Comparing the 'Then' with the 'Now', the text concludes with: "At last we are in a position to begin to reflect performance, and not our own competence, in our 21st century dictionary entries" (p. 271).

4.18   Duval's *Equivalence in Bilingual Dictionaries* (1991)

Initially published as an encyclopedia article in French, and translated into English by the author himself for PLR (an excellent translation at that), Duval's contribution is a welcome one indeed. Duval's topic is 'equivalence', which he extends to monolingual lexicography, where "there is equivalence between the entry word in the headword list from which the search starts and the body of the entry" (p. 273). From that angle, the title of his article is actually a misnomer. With reference to equivalence in bilingual lexicography — where the term is typically applied — Duval first points out that even so-called 'full equivalence' (between lemma sign and translation equivalent) does not always mean exact correspondence. He then proceeds by contrasting denotation vs. connotation, extension vs. comprehension, and language events vs. speech events. While the first dichotomy is also covered in OGPL (pp. 468-469 in there), the next two are not (and as such, constitute an informative extra).

Duval drew all his examples from the language pair French–English, two languages with very similar grammars. As a result, a whole range of additional thorny equivalence problems were avoided. A bilingual lexicographer working between a Bantu language and English, for example, is constantly faced with the problem that parts of speech in the one language do not correspond with those in the other (effectively turning, say, verbs into nouns, nouns into adjectives, etc.), or even with entire word classes (such as ideophones in Bantu) with no corresponding word classes nor translation equivalents (cf. e.g. De Schryver 2009). Bridging these various mismatches, in addition to the standard ones listed by Duval, is the real challenge of bilingual dictionary makers.

4.19   Church and Hanks's *Word Association Norms, Mutual Information, and Lexicography* (1989)

Church and Hanks's paper is the first of three in Part X, *On Tools for Lexicographers*. This paper had a galvanizing effect on the computational linguistics

world in 1989, when it was presented at the 27th annual meeting of the *Association for Computational Linguistics* (ACL) in Vancouver. There, it was the only paper to discuss statistical methods in computational linguistics, while at most (if not all) previous meetings of ACL there were none. Nowadays, such papers at ACL are in the majority.

The paper has attracted occasional hostile criticism, but seemingly only by people who feel threatened by it. For example, some people have proposed log-likelihood measures as a means of compensating for the so-called 'sparse data problem'. Computer scientists seem to like log-likelihood very much — it is elegant. But it has not been used in lexicography because it typically produces results that are less useful, practically speaking, than MI score (the statistical measure used by Church and Hanks in this paper) or t-score, which, as they were to comment in a later paper (Church et al. 1994) favours collocating function words, whereas MI favours collocations of pairs of content words.

The argument in the Church and Hanks paper is that collocations have a large role to play in decoding meaning, and that normal collocations are frequently recurrent in actual usage, so their relative importance can be measured by analysis of a large body of texts. What is more, Church and Hanks found (and published) a methodology for discovering the most significant collocates of any selected target word. The importance of this cannot be underestimated. Previous studies measured relations between two pre-selected target words, so they did not give us a *discovery* procedure. Church and Hanks then continue to show how collocates can be grouped to decide meaning.

When the proceedings of recent corpus linguistics conferences are read, it is surprising and saddening to note that there are many corpus linguists who have still, twenty years on, not yet adjusted their thinking to the most fundamental theoretical implication of this paper, namely that natural languages are analogical systems built around prototypes of many different sorts, and that corpora make it possible to identify these prototypes and measure agreement and variance statistically. If Church and Hanks are right about this (and their implication is hard to refute), it means that all linguistic categorization is a statistical procedure, a point of fundamental importance for lexicography of many different genres, as well as for theoretical and corpus linguistics.

MI is not really a "tool for lexicographers" (cf. the heading of Part X) — but it is the foundation on which one of the best corpus tools for lexicographers (the Sketch Engine) is based. This is discussed in the next section.

4.20   Kilgarriff, Rychlý, Smrž and Tugwell's *Sketch Engine* (2004)

The Sketch Engine is basically Manatee/Bonito to which word-sketch functionality was added. Manatee is a corpus query system (CQS), and Bonito its graphical user interface (GUI), both developed at Masaryk University, in Brno, by Pavel Rychlý (cf. e.g. Rychlý 2007). A word sketch is an automatically produced, corpus-based summary (i.e. 'sketch') of a word's grammatical and collo-

cational behaviour, first introduced by Adam Kilgariff and David Tugwell (2001). The Sketch Engine is arguably a magnificent tool for lexicographers (and corpus linguists in general), as it can be seen as a collocationally-annotated menu or index directly into the corpus. Figure 4, for instance, shows the word sketch for the noun *upset* in the Collins Wordbanks*Online*.

Home | Concordance | Word List | **Word Sketch** | Thesaurus | Sketch-Diff
Turn on clustering | More data | Less data | Save

upset   **WordBanks Online: English freq = 3866**   change options

object_of 1219 3.2		subject_of 364 1.5		a_modifier 1312 3.3		n_modifier 505 2.6		pp_over-p 22 6.0	
stomach	6 6.68	jumble	14 9.06	digestive	28 8.15	tummy	28 9.09	No.	4 2.05
cause	289 6.36	ruin	20 6.26	gastric	13 7.8	stomach	125 8.33		
spring	16 6.22	highlight	12 4.66	first-round	28 7.42	decathlon	7 8.17	**pp_in-p**	**183 3.2**
bar	13 5.6	occur	5 2.11	stunning	58 7.18	cup	11 3.65	history	39 3.4
score	48 4.86	cause	10 1.51	gastrointestinal	6 6.43	bit	40 3.14	tournament	8 2.84
pull	38 4.39	happen	6 0.92	third-round	5 6.05	surprise	4 1.89	round	14 1.91
avoid	22 3.77	continue	4 0.53	monumental	6 5.99	election	10 1.11	competition	4 0.96
stage	5 3.62			second-round	6 5.91	nothing	8 0.57	race	5 0.6
post	6 3.4	**adj_subject_of 56 1.9**		hormonal	5 5.87				
threaten	12 3.35	special	9 1.63	shocking	14 5.81	**modifies 164 0.3**		**pp_of-p**	**171 1.6**
suffer	16 2.84	possible	5 0.57	improbable	5 5.61	'Zona	24 11.99	No.	20 4.37
predict	5 2.67			emotional	45 5.5	No.	67 6.11	tournament	8 2.84
produce	19 2.55			major	150 5.3	tradition	4 1.82	century	9 1.67
create	17 2.13			big	243 4.87			season	20 1.66
record	4 2.11			electoral	9 4.7	**and/or**	**385 1.3**	decade	4 1.26
provide	14 1.37			minor	16 4.68	diarrhoea	10 6.84	round	6 0.69
get	72 1.36			mild	8 4.67	rash	5 5.35	day	17 0.2
include	20 1.21			mighty	4 4.63	nausea	4 5.34		
consider	5 0.58			Greek	11 4.62	headache	6 4.19		
seem	8 0.58			sensational	4 4.4	cold	6 4.11		
believe	5 0.27			historic	9 4.27	streak	8 3.92		
see	17 0.11			least	11 4.17	controversy	6 3.43		
				huge	37 4.1	anger	5 2.85		
				memorable	4 4.0	pain	7 1.93		

Figure 4: Word sketch for *upset* (noun) in the Collins Wordbanks*Online*

For each grammatical relation (object of, subject of, etc.), lists are presented with the words that typically combine with the search word (1st columns), in order of statistical significance (3rd columns). The hyperlinked numbers (2nd columns) stand for the number of concordance lines, and clicking on them reveals all and only those instances of this particular combination.

In addition to word sketches, the Sketch Engine also offers a corpus-based thesaurus, and sketch differences. Computationally, 'all' that is required to obtain all of this is a corpus of a particular language, as well as a lemmatizer, a POS-tagger and a (regular expression) grammar for that language. At present, the Sketch Engine is available for about a dozen languages.

Note that, although extremely powerful, the Sketch Engine is not the most widely used corpus tool. That honour goes to WordSmith. One of the reasons is probably that the Sketch Engine "is designed for use over the web, with a server holding the data" (p. 304). In contrast, WordSmith is a lightweight stand-alone application, which enables users to use their corpus tool even when off-line and to keep their valuable corpora to themselves, rather than to have their data stored on a server in the Czech Republic. Moreover, lexicographers who compile their dictionaries with the dictionary production system TshwaneLex typically do not use any other corpus tools, as TshwaneLex contains a built-in and seamlessly integrated CQS. That system handles both raw and POS-tagged corpora (cf. De Schryver and De Pauw 2007), with which examples may be transferred directly from the corpus to the appropriate senses, and provides the core concordance functionality expected from a CQS. Word-sketch-like features are, at present, not included.

4.21 Grefenstette's *The Future of Linguistics and Lexicographers: Will there be Lexicographers in the Year 3000?* (1998)

Grefenstette starts by pointing out that generative linguistics is doctrinaire, like communism (with different, mutually disparaging factions) — deviance from the party line is severely punished, while corpus linguistics allows laissez-faire, like unbridled capitalism. He then argues that a middle course is needed, which he terms 'approximate linguistics'. Actually, today's techniques in computational linguistics are exactly that, "approximate linguistics, i.e. approximations to linguistic theory that are both incrementally perfectible as well as being robust and immediately useful" (p. 320-321). His paper describes the state-of-the-art of approximate linguistics a decade ago — walking the reader through tokenization, morphological analysis and lemmatization, POS-tagging, shallow parsing, and semantic tagging — and "gives a hint of what approximate linguistics can offer working lexicographers today in their task of describing some fundamental sense of a word" (p. 321). His vision of the future too is intriguing (cf. De Schryver 2003: 163 *ff.*), and where that future has already arrived, he was right (cf. De Schryver 2003: 180).

Finally, with regard to the question he poses in the title, he answers that there will always be aspects of the lexicographic enterprise where humans will be needed (compare also Section 4.7 above). These aspects are not really spelled out, however, and for this reason this text, despite being full of thought-provoking visions, is of only marginal relevance to a reader on practical lexicography.

4.22 Miller G.A., Beckwith, Fellbaum, Gross and Miller K.J.'s *Introduction to WordNet: An On-line Lexical Database* (1990)

Miller et al.'s text is the only one in Part XI, *On Semantic Networks and Wordnets*.

WordNet is an online thesaurus in which lexical items are arranged into sets of synonyms, called 'synsets', with senses across synsets related through lexical relations such as hyponymy/hyperonymy, meronymy/holonymy, troponymy, antonymy, pertainymy, etc. WordNet does not feature anywhere in OGPL (apart from merely being mentioned in two places, on pp. 82 and 159), has nothing to do with the production of commercial MLDs (compare also Section 4.16 above), and has little relevance for practical lexicography. It is used in NLP research though — simply because it happens to be available.

4.23   Atkins and Varantola's *Monitoring Dictionary Use* (1997)

Atkins and Varantola's text is the only one in the last part, Part XII, *On Dictionary Use*. It is a meticulously described account of a meticulously planned and meticulously executed experiment in monitoring how people actually use their dictionaries, monolingual and bilingual. A paper version of the think-aloud protocol is used, whereby a first subject, the 'dictionary user', performs a translation task using one or more paper dictionaries of their choice, and a second subject, the 'recorder', keeps track of all the look-ups needed in an attempt to solve a particular problem (i.e. a search). A total of 103 students performed 574 searches involving exactly 1 000 look-ups. The details of each of the 103 students were written down on a so-called 'cover sheet', and the recorder recorded the details of each look-up on a so-called pre-prepared 'recorder sheet', one per look-up. All the collected material was keyed into a database, and then analysed with the aim to provide answers to the following questions:

— How do people consult dictionaries?
— What kind of information are users looking for?
— Where do users find the information they need?
— Do users find what they look up?
— Are people satisfied with what they get?
— What do people do when frustrated?
— When do people use an L2 monolingual dictionary?

Answers are indeed obtained for each of those questions, with one of the most impressive outcomes, according to Atkins and Varantola, "the amount of reassurance sought from their dictionaries, particularly about L2 collocation, by even the most skilled of non-native L2 speakers, however experienced in translation they may be" (p. 371).

The aim of Atkins and Varantola "was to monitor the dictionary look-up process in as natural a situation as possible" (p. 338), and for the time (the experiments were carried out in 1991 and 1993, with paper dictionaries only) their approach was indeed the state-of-the-art. Their study rightfully attracted many citations, as seen in Figure 3 (where it is actually the first one following

the NLP section). These days, electronic dictionaries are studied using the computer as an invisible and unobtrusive 'recorder', making the dictionary look-up process even more natural. Modern experiments range from the monitoring of the use of real dictionaries consulted in a natural setting (see for an early example, De Schryver and Joffe 2004), to test lexica consulted in an experimental set-up (see for the most recent study, Lew and Doroszewska 2009).

5.   How good is this Reader?

Bringing the quantitative and qualitative data together, one obtains Table 5.

Table 5: Contrasting the quantitative and qualitative evaluations of PLR

Ch.	Author(s)	Year	Citations	Refs from OGPL	Relevant to OGPL?
2	Johnson S	1747	–	***	***
3	Atkins BTS	1992/93	–	***	*
4	Apresjan JD	2002	–	**	**
5	Biber D	1993	**	**	–
6	Kilgarriff A & Grefenstette G	2003	**	**	–
7	Fillmore CJ	1992	–	**	**
8	Hanks P	2000	*	**	***
9	Kilgarriff A	1997	**	**	–
10	Stock PF	1984	–	***	***
11	Cowie AP	1994	–	**	–
12	Fontenelle T	1997	*	*	*
13	Bolinger D	1985	–	**	**
14	Rundell M	2006	–	*	**
15	Laufer B	1992	–	*	**
16	Rundell M	1998	*	***	***
17	Atkins BTS	2002	*	**	*
18	Duval A	1991	–	*	**
19	Church KW & Hanks P	1989	***	*	***
20	Kilgarriff A, Rychlý P, et al.	2004	**	**	**
21	Grefenstette G	1998	–	**	–
22	Miller GA, Beckwith R, et al.	1990	***	*	–
23	Atkins BTS & Varantola K	1997	*	**	***

In Table 5, the most important column is the last one, the qualitative evaluation, which is of course highly subjective. From it, and in the opinion of the present reviewer, one sees that 6 texts had better not been selected for PLR, while 3 are acceptable, 7 are good, and 6 are excellent choices. Contrasting this evaluation with the quantitative data, one notices a rather good overlap with the number of references from OGPL. A comparison of both of these with the number of citations is less convincing, however, but as was explained in Sec-

tion 3, (a) citation patterns show a bias towards papers in computational linguistics, and (b) in terms of authors, a much better fit is found.

Every reviewer of an anthology or reader of previously published papers no doubt has their own favourite authors and expectations, and the present reviewer is no exception. Among the selected papers are indeed some weak and/or less relevant ones — whose inclusion is surprising — while there are some even more surprising omissions. For example, it is quite astonishing that a modern reader on practical lexicography contains not a single paper by John Sinclair, the leading authority on collocations and the founder of statistical approaches to corpus linguistics, although papers by his colleagues (Stock, Hanks) are included. Collocations and phraseology are represented only by a single short encyclopedia article by Cowie, which reflects attitudes of the pre-corpus era and now seems rather out of date. This omission is perhaps not an oversight: It might to be the result of rivalry between John Sinclair and Sue Atkins (co-author of OGPL, which the volume under review is designed to support). It perhaps also explains why Hanks's important paper on 'Definitions and Explanations' is not included. The COBUILD full-sentence style of definition is represented only by a couple of less than enthusiastic accounts authored by Michael Rundell, the editor of a rival dictionary (MEDAL). The same reason may underlie the absence of Fox's 'The Case for Examples', another important product of the COBUILD project.

It is good to see the Russian tradition represented by Apresjan, but the equally important Mel'čuk appears only by proxy, in an article by Fontenelle himself. The generative tradition is not represented: One might have expected a paper each by Jackendoff, Pustejovsky, and Lakoff. Based on Table 3, Bogaards, Geeraerts, Hoey, and Moon could and should also have been included. For good measure, a paper by Wierzbicka ought to have balanced out PLR.

In conclusion, this is a valuable reader of some very good papers and a handful of less good ones. However, its lack of balance means that supplementary reading is necessary for any would-be lexicographer.

Acknowledgements

Endnotes

1.   While elegant and space-saving, the disadvantages are that one cannot, from the references, quickly see what a particular paper is about or who is and is not quoted, while one will always need the entire book to keep things complete.

2.   Given that the various papers which currently get zero hits in Google Scholar have most definitely been cited as well, one may be led to conclude that Google Scholar is particularly unreliable as an index of lexicographic scholarly activity. This is partly because it does not

survey all published sources, but only certain ones. If *all* publications were surveyed and if the time frame were 250 years, Johnson (1747), it would seem, would surely come out way ahead of Church and Hanks (1989).

3. Fontenelle could have considered selections, however, as Hartmann did for his anthology (cf. Table 1).

4. Unless otherwise noted, all page numbers preceded by 'p.' or 'pp.' are with reference to Fontenelle's PLR.

5. Actually, Fontenelle's PLR itself is now also and already available through Google Books, both in snippet view and as a limited preview.

6. Note that the collocational preferences of English *arrive* have changed since Johnson's time.

7. The following paragraph is taken from De Schryver (2005a: 422), where Apresjan's contribution was summarized.

8. Although both volumes have gone through a second edition (cf. the references), a thorough and truly corpus-driven revision is still outstanding. One of the selling points that accompanies the 2nd edition of volume 1 reads: "Examples of use taken from a wide range of contemporary sources, many drawn from the Oxford Corpus of the English Language" (http://www.oup.com/elt/catalogue/isbn/6614?cc=global), and for the 2nd edition of volume 2, tellingly shorter: "Example sentences taken from a wide range of contemporary sources" (http://www.oup.com/elt/catalogue/isbn/6612?cc=global). Selecting examples to support what was already there is very different from compiling such a dictionary from scratch driven by corpus evidence.

References

Atkins, B.T.S. and M. Rundell. 2008. *The Oxford Guide to Practical Lexicography.* New York: Oxford University Press.

Bogaards, P. 2009. EURALEX Newsletter — Autumn 2009. *International Journal of Lexicography* 22(3): 351-360.

Church, K.W., W. Gale, P. Hanks, D. Hindle and R. Moon. 1994. Lexical Substitutability. Atkins, B.T.S. and A. Zampolli (Eds.). 1994. *Computational Approaches to the Lexicon*: 153-177. Oxford: Clarendon Press.

Collins Wordbanks*Online*. 2009. An online service which gives subscribers access to language data based on the Collins corpora of modern written and spoken text [online]. http://wordbanks.harpercollins.co.uk/Docs/Help/guide.html

Cowie, A.P. and R. Mackin (Eds.). 1975. *Oxford Dictionary of Current Idiomatic English. Volume 1: Verbs with Prepositions and Particles* (Retitled *Oxford Dictionary of Phrasal Verbs* for the second edition, 1993). Oxford: Oxford University Press.

Cowie, A.P., R. Mackin and I.R. McCaig (Eds.). 1983. *Oxford Dictionary of Current Idiomatic English. Volume 2: Phrase, Clause and Sentence Idioms* (Retitled *Oxford Dictionary of English Idioms* for the second edition, 1993). Oxford: Oxford University Press.

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. Book Review: R.R.K. Hartmann, ed. 2003. *Lexicography: Critical Concepts* (Routledge, three volumes, 1 327 pages). *International Journal of Lexicography* 18(1): 93-102.

De Schryver, G.-M. 2005a. Book Review: M.-H. Corréard, ed. 2002. *Lexicography and Natural Language Processing. A Festschrift in Honour of B.T.S. Atkins. Lexicographica: International Annual for Lexicography* 21: 420-425.

De Schryver, G.-M. 2008. An Analysis of *The Oxford Guide to Practical Lexicography* (Atkins and Rundell 2008). *Lexikos* 18: 423-445.

De Schryver, G.-M. 2008a. Book Review: P. Hanks, ed. 2008. *Lexicology: Critical Concepts* (Routledge, six volumes, 2 793 pages). *International Journal of Lexicography* 21(4): 419-438.

De Schryver, G.-M. 2009. The Lexicographic Treatment of Ideophones in Zulu. *Lexikos* 19: 34-54.

De Schryver, G.-M. 2009a. Bibliometrics in Lexicography. *International Journal of Lexicography* 22(4).

De Schryver, G.-M. and G. De Pauw. 2007. Dictionary Writing System (DWS) + Corpus Query Package (CQP): The Case of *TshwaneLex. Lexikos* 17: 226-246.

De Schryver, G.-M. and D. Joffe. 2004. On How Electronic Dictionaries are Really Used. Williams, G. and S. Vessier (Eds.). 2004. *Proceedings of the Eleventh EURALEX International Congress, EURALEX 2004, Lorient, France, July 6–10, 2004*: 187-196. Lorient: Faculté des Lettres et des Sciences Humaines, Université de Bretagne Sud.

Fillmore, C.J. and B.T.S. Atkins. 1992. Toward a Frame-based Lexicon: The Semantics of RISK and its Neighbors. Lehrer, A. and E.F. Kittay (Eds.). 1992. *Frames, Fields and Contrasts: New Essays in Semantic and Lexical Organization*: 75-102. Hillsdale: Lawrence Erlbaum.

Fontenelle, T. (Ed.). 2008. *Practical Lexicography: A Reader*. New York: Oxford University Press.

Fox, G. 1987. The Case for Examples. Sinclair, J.M. (Ed.). 1987: 137-149.

FrameNet. 2009. An online lexical resource for English, based on Frame Semantics and supported by corpus evidence [online]. http://framenet.icsi.berkeley.edu/

Google Books. 2009. A search and preview of millions of books from libraries and publishers worldwide [online]. http://books.google.com/

Google Scholar. 2009. A search of scholarly literature across many disciplines and sources, including theses, books, abstracts and articles [online]. http://scholar.google.com/

Hanks, P. 1987. Definitions and Explanations. Sinclair, J.M. (Ed.). 1987: 116-136.

Hanks, P. 2005. Johnson and Modern Lexicography. *International Journal of Lexicography* 18(2): 243-266.

Hanks, P. (Ed.). 2008. *Lexicology: Critical Concepts* (6 volumes). Abingdon: Routledge.

Hanks, P. 2008a. The Lexicographical Legacy of John Sinclair. *International Journal of Lexicography* 21(3): 219-229.

Hanks, P. 2009. Book Review: S.J. Perrault, ed. 2008. *Merriam-Webster's Advanced Learner's English Dictionary. International Journal of Lexicography* 22(3): 301-315.

Hartmann, R.R.K. (Ed.). 2003. *Lexicography: Critical Concepts* (3 volumes). London: Routledge.

Johnson, S. 1755. *Johnson, Preface to the Dictionary.* Edited by Jack Lynch [online]. http://andromeda.rutgers.edu/~jlynch/Texts/preface.html

Jones, S. and J.M. Sinclair. 1974. English Lexical Collocations. A Study in Computational Linguistics. *Cahiers de Lexicologie* 24(1): 15-61.

Kilgarriff, A. 2009. *Adam Kilgarriff's Home Page* [online]. http://www.kilgarriff.co.uk/

Kilgarriff, A. and D. Tugwell. 2001. WORD SKETCH: Extraction and Display of Significant Collocations for Lexicography. *Proceedings of the ACL Workshop on 'COLLOCATION: Computational Extraction, Analysis and Exploitation'*: 32-38. Toulouse. [Reprinted as 'Sketching Words' in Cor-

réard, M.-H. (Ed.). 2002. *Lexicography and Natural Language Processing. A Festschrift in Honour of B.T.S. Atkins*: 125-137. Euralex.]

Lew, R. and J. Doroszewska. 2009. Electronic Dictionary Entries with Animated Pictures: Lookup Preferences and Word Retention. *International Journal of Lexicography* 22(3): 239-257.

Rundell, M. 2002. *Macmillan English Dictionary for Advanced Learners*. Oxford: Macmillan. [MEDAL]

Rychlý, P. 2007. Manatee/Bonito — A Modular Corpus Manager. Sojka, P. and A. Horák (Eds.). 2007. *Proceedings of Recent Advances in Slavonic Natural Language Processing, RASLAN 2007*: 65-70. Brno: Masaryk University.

Scott, M. 2009. *WordSmith Tools* [online]. http://www.lexically.net/wordsmith/

Senseval. 2009. Evaluation Exercises for the Semantic Analysis of Text [online]. http://www.senseval.org/

Sinclair, J.M. 1984. Naturalness in Language. Aarts, J. and W. Meijs (Eds.). 1984. *Corpus Linguistics. Recent Developments in the Use of Computer Corpora in English Language Research*: 203-210. Amsterdam: Rodopi.

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*. London: Collins ELT.

Sinclair, J.M. and G. Fox. 1995. *Collins COBUILD English Dictionary*. London: HarperCollins. [COBUILD-3]

Sinclair, J.M. and P. Hanks. 1987. *Collins COBUILD English Language Dictionary*. London: Harper-Collins. [COBUILD-1]

TshwaneLex. 2009. TshwaneLex Dictionary Production System [online]. http://tshwanedje.com/tshwanelex/

Addendum: Fact sheet of the 23 texts in *Practical Lexicography: A Reader*

Author(s)	Born	Year	Title	Source	P.	Ch.	Pp.	Online	Cited
Apresjan JD	1930	2002	Principles of Systematic Lexicography	Fest - Atkins	I	4	10	–	–
Atkins BTS	1931	1992/93	Theoretical Lexicography and its Relation to Dictionary-Making	J - Ds: DSNA	I	3	20	–	–
Atkins BTS	1931	2002	Then and Now: Competence and Performance in 35 Years of Lexicography	Proc - Euralex '02	IX	17	26	–	5
Atkins BTS & Varantola K	1931, 1946	1997	Monitoring Dictionary Use	J - IJL	XII	23	39	$	48
Biber D	(1946)	1993	Representativeness in Corpus Design	J - Ling & Lit Comp	II	5	25	$	261
Bolinger D	1907	1985	Defining the Indefinable	Ed coll	VI	13	4	–	–
Church KW & Hanks P	(1957), 1940	1989	Word Association Norms, Mutual Information, and Lexicography	Proc - ACL	X	19	11	I	1568
Cowie AP	1931	1994	Phraseology	Encyclopedia	V	11	5	–	–
Duval A (Tr by the author)	(1949)	1991	Equivalence in Bilingual Dictionaries	Encyclopedia	IX	18	10	–	–
Fillmore CJ	1929	1992	'Corpus Linguistics' or 'Computer-aided Armchair Linguistics'	Proc - Nobel '82	III	7	18	G	–
Fontenelle T	1964	1997	Using a Bilingual Dictionary to Create Semantic Networks	J - IJL	V	12	21	$	12
Fontenelle T	1964	2008	Editor's Introduction	N.A.	0	1	15	N.A.	N.A.
Grefenstette G	1956	1998	The Future of Linguistics and Lexicographers: Will there be Lexicographers in the Year 3000?	Proc - Euralex '98	X	21	17	–	–
Hanks P	1940	2000	Do Word Meanings Exist?	J - Comp & Hum	IV	8	10	(I)	47

Author			Title						Pp.
Johnson S (Ed by Jack Lynch)	1709	1747	The Plan of a Dictionary of the English Language	Letter	I	2	12	I	–
Kilgarriff A	1960	1997	I Don't Believe in Word Senses	J - Comp & Hum	IV	9	17	I	156
Kilgarriff A & Grefenstette G	1960, 1956	2003	Introduction to the Special Issue on the Web as Corpus	J - Comp Ling	II	6	13	I	299
Kilgarriff A, Rychlý P, Smrž P & Tugwell D	1960, 1973, 1972, ?	2004	The Sketch Engine	Proc - Euralex '04	X	20	10	I	142
Laufer B	?	1992	Corpus-based versus Lexicographer Examples in Comprehension and Production of New Words	Proc - Euralex '92	VII	15	6	–	–
Miller GA, Beckwith R, Fellbaum C, Gross D & Miller KJ	1920, ?, (1950), 1968, ?	1990	Introduction to WordNet: An On-line Lexical Database	J - IJL	XI	22	8	$	1386
Rundell M	1950	1998	Recent Trends in English Pedagogical Lexicography	J - IJL	VIII	16	23	$	32
Rundell M	1950	2006	More Than One Way to Skin a Cat: Why Full-Sentence Definitions Have Not Been Universally Adopted	Proc - Euralex '06	VI	14	13	–	–
Stock PF	(1949)	1984	Polysemy	Proc - Euralex '83	IV	10	8	–	–

Abbreviations: Ed = edited, Tr = translated; P. = Part, Ch. = Chapter, Pp. = pages; $ = by subscription, G = Google Books, I = Internet, (I) = idem but different version

Notes: One date was adapted here (Church & Hanks 1990 → 1989) compared to the one given in *Practical Lexicography: A Reader*, to better reflect the *original* year of publication. Years between brackets are informed extrapolations.