

On How Electronic Dictionaries are Really Used

Gilles-Maurice de Schryver^{#,°}, David Joffe[°]

Department of African Languages and Cultures, Ghent University, Rozier 44, 9000 Ghent, Belgium[#] (gillesmaurice.deschryver@UGent.be) || TshwaneDJe Human Language Technology, PO Box 299, Wapadrand 0050, South Africa[°] (david.joffe@tshwanedje.com)

Abstract

In this paper it is shown how a real-world electronic dictionary can be simultaneously compiled and its use studied. While the results of the dictionary use study may be successfully fed back into the compilation, the semi-automatic analysis of the use itself for the first time reveals how electronic dictionaries are really used. In order to achieve this, an intricate and multifaceted integrated log file tracks every single action of every single user – date and time stamping each lookup, ordering founds and not-found, monitoring long-term vocabulary retention, etc. – with a multitude of summaries being presented to the lexicographers. The ultimate goal is that with such data the parameters of various user profiles could be pinpointed, with which self-tailoring electronic dictionaries could be built.

1. Introduction: Research on (Electronic) Dictionary Use

Within the framework of Wiegand's dictionary research (*Wörterbuchforschung*), research on dictionary use (*Wörterbuchbenutzungsforschung*) is the first of four sub-fields in metalexicography. In one of his monographs, Wiegand (1998) devotes nearly a thousand pages to the issue, citing 86 previous studies on the way. In the same year, Hulstijn & Atkins (1998) list some 50 papers revolving around empirical research into dictionary use; their study itself being part of an edited collection entitled 'Using Dictionaries' (Atkins 1998). Around a dozen brief summaries of articles on dictionary usage may also be found in Dolezal & McCreary (1999). Mainly in the more recent contributions, the virtues of computer-controlled investigations are emphasised, as these enable an unobtrusive monitoring of dictionary use. With the popularisation of electronic dictionaries, an increasing number of scholars have pointed out the potential of especially Internet dictionaries to generate free implicit feedback, retrievable from log files attached to such dictionaries (De Schryver 2003a: 160-161).

Although the proposal to draw upon log files in order to *improve* dictionaries was already expressed in the mid-1980s (Abate 1985; Crystal 1986), and although numerous researchers have reiterated this idea in recent years (Hulstijn & Atkins 1998; Sobkowiak 1999; Docherty 2000; Harley 2000; Sato 2000), very few reports have been published of *real-world* dictionaries actually making use of this strategy. Notable exceptions are Löfberg (2002) and Prószyński & Kis (2002). Instead, electronic dictionaries cum log files seem to be more popular in *research environments* focusing on vocabulary acquisition (Hulstijn 1993; Knight 1994; Hulstijn & Trompeter 1998; Laufer 2000; Laufer & Hill 2000). When it comes to electronic dictionaries, statements regarding log files are often hypothetical, such as in: 'A log file of user access and queries is kept that *should* serve to give insight on how such a service is used' (Popescu-Belis et al. 2002: 1144 [emphasis added]). What is true for log files, is also true for the utilisation of direct feedback, whereby users are encouraged to comment online on

dictionary articles and to suggest new items (Dodd 1989; Carr 1997; Considine 1998; Harley 2000; Nesi 2000; Warburton 2000), i.e. reports on what is done with this type of feedback are hard to come by.

2. Simultaneous Feedback (SF) and Fuzzy SF

Since 1997 a methodology is being worked out whereby the results of studies of actual dictionary use are directly integrated into the compilation of a reference work (De Schryver 1999). Feedback from the envisaged target user group is systematically and continuously obtained while compilation is still in progress. In practical terms this process, known as *Simultaneous Feedback* (SF), ‘can be understood as entailing a method in terms of which the release of several small-scale parallel dictionaries triggers off feedback that is instantly channelled back into the compilation process of a main dictionary’ (De Schryver & Prinsloo 2000: 197).

Whereas the original concept was developed within the confines of printed dictionaries, work on an electronic adaptation, known as *Fuzzy SF*, was begun in 2001 (De Schryver & Prinsloo 2001). In *Fuzzy SF*, traditional means for gathering feedback such as participant observation or questionnaires are replaced with the computational tracking of all actions in an electronic dictionary. Ultimately, the idea is that an automated analysis of the log files will enable the dictionary to tailor itself to each and every particular user. At present, the analysis of the log files is still largely done manually, in part with the aim to draw up typical user profiles that will then be fed into the projected adaptive and intelligent dictionary of the future (cf. De Schryver 2003a: 188-190).

3. The Sesotho sa Leboa Dictionary Project (SeDiPro)

One electronic reference work that is currently being compiled within the framework of *Fuzzy SF* is the *Sesotho sa Leboa Dictionary Project* (SeDiPro), a bilingual dictionary between Sesotho sa Leboa (a Bantu language spoken in South Africa) and English. This dictionary has been made available on the Internet (<http://africanlanguages.com/sdp/>), and with approximately 25 000 items on the Sesotho sa Leboa side and 28 000 in the English search index, as well as a linguistics terminology list of over 300 items, it is currently the largest freely-available African-language Internet dictionary. It is also the first African-language Internet dictionary allowing for some low-level fuzzy searches and for which the entire interface can be set to an African language in addition to English. The latter is not merely a small extra or a political move, but a fully functional component. Indeed, ‘the terminology list contains a world’s first for an online dictionary, namely the customisation of the output of part-of-speech (POS) tags, usage labels and cross-references depending on the language chosen’ (De Schryver 2003b: 12). This customisation is realised in real time on the Internet.

Unlike in the great majority of current electronic dictionaries, dealing with feedback, whether implicit or explicit, has been a central component right from the start in SeDiPro. Since the first day the dictionary was posted online, a well thought out log file has been unobtrusively keeping track of all aspects of dictionary use, while an online feedback form has allowed for a more traditional and open way of receiving feedback.

The software used is *TshwaneLex* (see Joffe & De Schryver 2004), a modern dictionary compilation program developed by *TshwaneDJe HLT* (<http://tshwanedje.com/>). Using *TshwaneLex*, the dictionary contents are exported to a MySQL (<http://www.mysql.com/>) database that is stored on a Linux/Apache web server (<http://www.apache.org/>). The online dictionary software is implemented using the scripting language PHP (<http://www.php.net/>). The PHP scripts generate the HTML-based search interface, log users' searches, query the MySQL database and generate the search results. In order to track the number of unique visitors (and to track return visits), each visitor is associated with a unique 128-bit identification number (visitor ID) that is saved in the log file each time that user performs a search. The visitor ID is generated the first time a user uses the dictionary, and it is stored using a cookie in the user's web browser software, allowing the ID to be retrieved again during subsequent searches or visits. This method, which is commonly used on the Internet, has the advantage of being unobtrusive and transparent to the user (as opposed to, for example, a user registration system with login and password, as is the case for *ELDIT*). However, it is not foolproof. Some users periodically clear the cookies stored by their browser, and other users even block cookies altogether. One can also not distinguish between multiple users who share a computer, or determine when a single user has made use of multiple computers (e.g., a student who uses a computer lab). Nonetheless, the technique is reliable in the majority of cases, providing an error margin of probably not more than 15%.

4. Analysing the SeDiPro Log Files

SeDiPro was made available on 22 April 2003, and what follows is a brief analysis of the main dictionary log files of the first six months, referred to as months 1 to 6 below. A total of 21 337 lookups were made by 2 530 different visitors during those six months. Although this corresponds to an average of 8.4 lookups per visitor, the actual distribution is Zipfian, as shown in Figure 1, with the great majority of users only performing a few searches each, and a minority up to several hundreds of searches each.

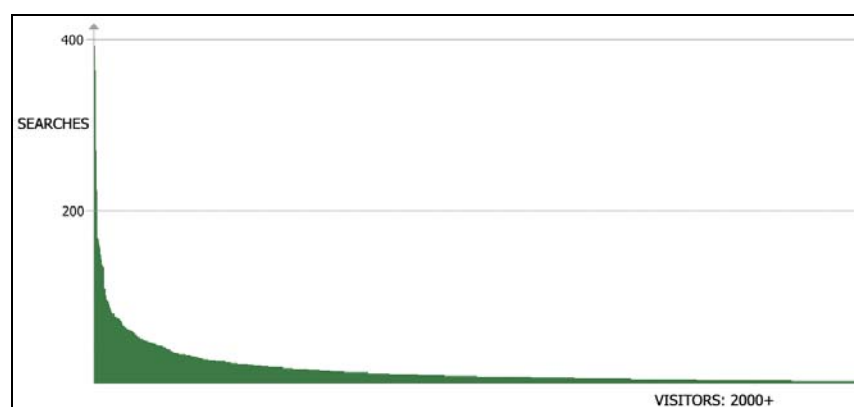


Figure 1: Distribution of the number of searches per visitor

Per day, an average of 16.8 visitors carried out 116.6 lookups: 40.8 (i.e. 35%) in the direction Sesotho sa Leboa to English, and 75.8 (i.e. 65%) in the reverse direction. Lookups during weekends were markedly less frequent than during weekdays. With each passing month, both the number of searches and the number of visitors increased, as can be seen in Figure 2.

Monthly				
Month	Total searches	Sesotho sa Leboa	English	Unique visitors
1	1308	512	796	138
2	2261	915	1346	282
3	5892	2266	3626	851
4	3607	1109	2498	482
5	4596	1647	2949	529
6	3673	1025	2648	427

Figure 2: Distributions of the number of searches and visitors per month

The peak for month 3 is mainly the effect of a series of local media releases in several South African newspapers starting on 3-4 July 2003, and lasting for a week. In contrast, the official academic launch two weeks earlier, on 20 June 2003, had no great impact on the stats. When one studies the origin of all searches per domain or country one may conclude that the large majority of the lookups are also made from within South Africa.

Apart from such general facts about electronic dictionary use, the logs also enable one to examine more closely what people actually look up. This could for example be linked to the following research question: ‘Are the top 100 searches also the top 100 in a corpus?’ If it would turn out that there is indeed a large overlap, this finding would provide substantial support for the practice of including or omitting lemma signs in a dictionary based on frequency considerations (and by extension for corpus-based lexicography in general).

If one compares the top 100 Sesotho sa Leboa searches with the ranks of the corresponding items in a frequency list derived from a 6.1-million-word Sesotho sa Leboa corpus, then one notices that 30 of the top 100 searches can also be found in the corpus top 100, while as many as 63 can be found in the corpus top 1 000. Clearly, users indeed look up the frequent words of the language. In the top 100 searches there are a further 6 foreign words (4 Setswana and 2 English), and of the remaining 31 words no less than 17 either have to do with the sexual sphere or are extremely offensive: *marêê* ‘testicles’, *masepa* ‘(off.) shit’, *mogwêê* ‘(off.) anus’, *mpopo* ‘(off.) private part (vagina; penis)’, *nnyô* ‘vagina’, *nnywana* ‘(off.) cunt’, *ntoto* ‘penis’, *nyôba* ‘(vulgar) fuck’, *sefêbê* ‘prostitute; (off.) bitch’, *thôbalanô* ‘sex’, etc. This latter phenomenon might very well be the case for all (Internet) dictionaries. The following note was for example added to the yearly top 50 of the *Cambridge Dictionaries Online* (<http://dictionary.cambridge.org/>): ‘The list had to be edited slightly for prurient content’. One may therefore conclude that genuine frequent words are looked up on the one hand, and then those words that only mother-tongue speakers know but, as they are taboo, *never* pronounce in public.

An analogous study of the top 100 English searches reveals a similar pattern, with 18 of the top 100 searches also in the BNC top 100 (Leech et al. 2001) and 62 in the BNC top 1 000. A single item in the top 100 searches is misspelled, while 6 of the remaining 37 searches again belong to the same sexual/offensive sphere: *bitch*, *fuck*, *penis*, *sex*, *shit* and *vagina*.

The most frequently looked for word in both sides of the dictionary is the same concept: *hello* and *dumêla* respectively. As a matter of fact, numerous users seem to ‘greet the dictionary’ on arrival (other frequent variants are (*good*) *morning* and *thôbêla* respectively), and after having gone through their lookups, they also ‘say’ *goodbye* and *sepela* (*gabotse*) respectively.

Not all lookups are successful, of course. Quite a number of English words are searched for in the Sesotho sa Leboa side and vice versa. A surprisingly high percentage of words are also consistently misspelled and/or mistyped in both languages, and a large number of items from neighbouring African languages (especially from Setswana and isiZulu), as well as from Afrikaans and Dutch, are found in the logs. Not-founds that are repeatedly searched for and that should have been included in the first place, are added during updates of the online dictionary. Care is taken not to react blindly on incidental peaks however. Compare in this regard the sudden popularity of so-called terrorist-related concepts following September 11, 2001, as recorded and analysed by for instance *Google* (<http://www.google.com/press/zeitgeist/9-11.html>) and *Cambridge Dictionaries Online* (http://dictionary.cambridge.org/top20/top20_0901.asp).

Two substantial updates of SeDiPro were effected, the first on 30 July 2003, the second on 4 September 2003. Although no more than a dozen, respectively around one hundred, changes/additions were made, the impact on the hit rate was impressive. The percentage of successful searches in the English side, for example, went from 67% to 72% after the first update, and then to 75% after the second update. Examples from the first update include the addition of ‘fuck’ as translation equivalent of *nyôba*, to supplement the more formal ‘have sexual intercourse’, or the inclusion of *Inthanêê* ‘Internet’ which had not yet been lemmatised. During the second update more abstract concepts as well as multi-word units (MWUs) were included, for example ‘devaluation’ was added as an equivalent of *phôkôtsô*, while MWUs such as *ka kua ga* ‘across’ or *ya/tša/wa/... gauta* ‘golden’ were lemmatised. The latter necessitated the creation of a new convention; in other instances new words had to be coined following consultation and fieldwork. As such, the online dictionary could be viewed as a *service* to the community.

On another level, especially relevant within the framework of Fuzzy SF, one can zoom in on particular users, and study their search strategies. Research questions may include: ‘How do users look up?’ and ‘What do users do when a search fails?’ As the log files are written in such a way that they automatically track and summarise all aspects of each and every visitor, one may for example study look-up behaviour when the exact spelling is not known. One series of searches is: ... chinese ✓ ¶ indian ✓ ¶ arab ✓ ¶ foriegner × ¶ forienger × ¶ outsider ✓ ¶ foreigner ✓ ... Here the user struggled with the spelling of ‘foreigner’, then tried the related word ‘outsider’, which returned the article for *lephatle*. As the latter article also mentions ‘foreigner’ in its translation equivalent paradigm, the user now saw the correct spelling, and went on to search for it directly. The beginning of the above series is also rather typical in the logs, as items from a single onomasiological field are looked up. In some cases, users even type in single search strings such as ‘eye,ear,nose,mouth,tongue,hand,arm’.

The online dictionary has several dozen regular visitors. In Figure 3 the searches made by one such visitor, between 19 June and 12 September 2003, are shown. During the studied period this visitor performed 168 searches, looked up in both Sesotho sa Leboa and English, and did not use the dictionary on weekends (grey). *Hlôgô* (automatically re-routed from *hlogo*) ‘head; prefix; heading; principal’ being rather polysemous, it is not surprising it was looked for repeatedly; yet the two searches for ‘woodpecker’ suggest that there was no long-term retention for the ways to express this item in Sesotho sa Leboa.

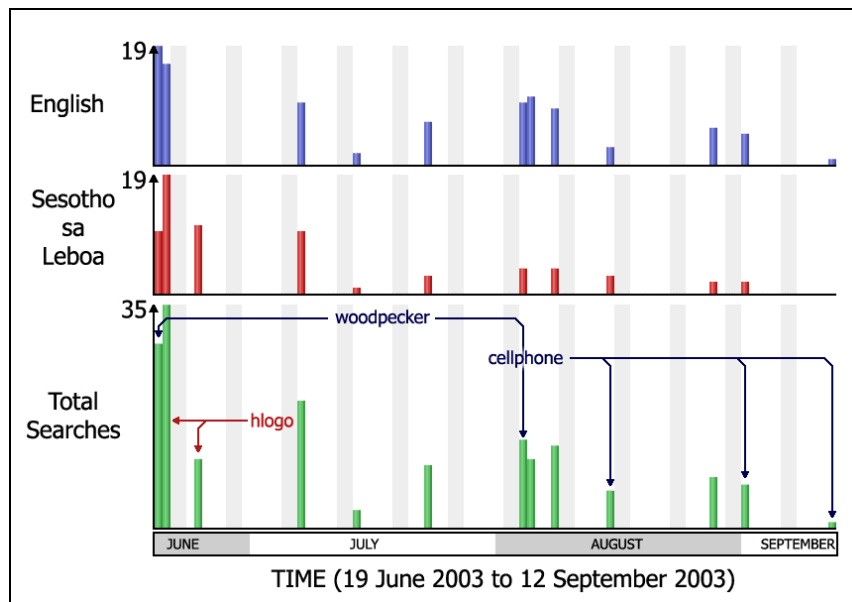


Figure 3: Zooming in on the searches across time by one particular visitor

5. Analysing the SeDiPro Online Feedback Forms

In addition to the unobtrusive logging of the dictionary users’ actions, as briefly discussed in the previous section, users are invited to provide the compilers with direct feedback. To that end, an online dictionary feedback form (offered in both Sesotho sa Leboa and English) is available, and all this feedback is automatically date stamped and logged. A total of 52 messages were received during the first 6 months, or thus on average two per week. In Table 1 the topics of these messages, as well as where the authors come from and the languages in which they write, are summarised.

Topic [N = 52]	%	Affiliation [N = 52]	%
Support for the project		Expected	
Congratulatory	31.7	Academic (.ac) / Education (.edu)	38.5
Expressing disbelief	8.7	Less expected	
		Company (.com) / Organisation (.org)	32.7
Regarding SeDiPro itself		Surfer	21.2
General contents	11.5	Professional translator	5.8
Useful tips	10.6	Government employee	1.9
Questions on specific word use	7.7		100.1
Extra grammar / synonyms	3.8		
please			
Interface	5.8		
In other dictionary media as well		Language [N = 52]	%
please		English	87.5
Electronic dictionary (on CD etc.)	8.7	Sesotho sa Leboa	6.7
Paper dictionary	3.8	Afrikaans / Dutch	5.8
More languages please	7.7		100.0
	100.0		

Table 1: Summary of the analysis of the online feedback forms

Four out of every ten messages is written in support of the project. In some cases the users even express outright disbelief, such as in ‘OH MY GOODNESS. Very incredible that this is indeed available! Please continue improving upon this wonderful work. Most of all, thank you for even beginning this project!!!! Many thanks!!!!’ or ‘Thank you a million – never would I have thought I’d see my language first, translated to English and secondly, on-line. This is remarkable and I hope this site is exposed to a lot of “ba gešo” [our people] out there! Re a leboga ka kudu! [We thank you very much!]’. Apart from being gratifying, it is important to realise that, in a country officially eleven-lingual, so far the majority of the people have just been paying lip service to all things relating to multilingualism. Another four out of every ten messages deals with the dictionary itself. On the one hand there are general remarks and useful tips, on the other there are questions that deal with specific words or requests for more data categories. An example of the first is ‘I strongly believe that the Northern Sotho word *seriti* has more to do with dignity or aura than just personality’, and of the second ‘Can I ask you for a translation of the word “cellphone” ...’. Surprisingly, 5.8% of the messages deal with aspects of the interface, with people commenting on computational features. In 12.5% of the messages people express their wish to acquire a stand-alone electronic and/or a hardcopy version. Lastly, in 7.7% of the messages users point out that they would like to see analogous dictionaries for the other South African languages, and request that these be compiled.

Each and every message is followed up, a reply is sent in all cases (which sometimes results in real discussion), and the dictionary is adapted whenever necessary. In line with the idea to be a service to the community, attending to the formal feedback receives higher priority

than implementing changes prompted by trends seen in the log files. Users also take the compilers – and with this perhaps even the concept of Simultaneous Feedback – seriously. In Figure 3, the first search for ‘cellphone’ (on August 15) was unsuccessful. During this user’s next visit (on August 28), the feedback form was filled in with a request that the compilers come up with a translation equivalent for this word. Already four days later (on September 1) this user checked again, yet the matter was still being researched. Three days later an acceptable translation (*mogalathêkêng*) was found and uploaded. A third search for ‘cellphone’ (on September 12) then returned a hit. As far as the call for Internet dictionaries in other South African languages is concerned, planning of ZuDiPro, the isiZulu equivalent of SeDiPro, has begun. On the whole one observes a very good correspondence between the formal feedback received through the online feedback forms, and the informal feedback obtained by means of an analysis of the log files. This is a satisfying sign indeed and indicates that modifying and adapting dictionary contents based on log-stat trends is a feasible strategy.

When one focuses on the origin of the feedback forms, one sees that 38.5% originate from the academic and higher education spheres. One would have expected this percentage to be even higher, given that 82% of all African-language Internet dictionaries are provided (and mainly used) by tertiary institutions, with only 13% by dotcoms and another 5% being private efforts (De Schryver 2003b: 9). The large percentage of visitors from dotcoms, organisations and individual surfers is thus surprising. Lastly, 87.5% of the messages are written in English, with only 6.7% in Sesotho sa Leboa and another 5.8% in Afrikaans or Dutch. One would have wished the African-language part to be larger.

6. Conclusion: Truly Unobtrusively Improving Dictionary Contents and Studying Dictionary Use

A *real* electronic dictionary used in a *natural* setting with *no manipulation* of research variables whatsoever was presented in this paper. It was shown how a combination of informal log-file analysis and the processing of formal online feedback forms may lead to improving dictionary contents and to a study of (electronic) dictionary use in a *truly unobtrusive* way.

With specific reference to a Sesotho sa Leboa Internet dictionary, it was indicated that the general trend during the first six months has been one of a growing number of lookups by a growing number of users – already reaching unexpected heights for an African language. Most visits remain local, and being mentioned in the popular press, unlike an academic launch, results in huge spikes in the stats. While the distribution of the number of lookups per visitor is Zipfian, most visitors tend to look up frequent items on the one hand, and sexual/offensive items on the other. Revisions and improvements of the dictionary may successfully be based on a semi-automatic analysis of log files, in combination with follow-ups on feedback received electronically. With the discussed tracking function, any number of individual users’ look-up strategies may be monitored across time, which is especially relevant for studying vocabulary retention and for drawing up user profiles needed for the projected intelligent and adaptive Fuzzy SF dictionary of the future.

More basic plans include smart re-routers for frequently misspelled and/or orthographically challenging words, whereby the software would not merely re-route but also point out the correct spelling. The selection of items to re-route ought to be based on frequencies derived from the log files, and in the case of frequently looked up foreign words a link to a dictionary in that particular language could also be offered. More than ever, dictionaries are becoming a *service* to the community.

References

- Abate, F. R.** 1985. Dictionaries Past & Future: Issues and Prospects. *Dictionaries* 7: 270–283.
- Atkins, B. T. S.** (ed.). 1998. *Using Dictionaries: Studies of Dictionary Use by Language Learners and Translators*. Tübingen: Max Niemeyer Verlag.
- Carr, M.** 1997. Internet Dictionaries and Lexicography. *International Journal of Lexicography* 10.3: 209–230.
- Considine, J. P.** 1998. ‘Why do large historical dictionaries give so much pleasure to their owners and users?’ in T. Fontenelle et al. (eds.), *ACTES EURALEX’98 PROCEEDINGS: 579–587*. Liège: English and Dutch Departments, University of Liège.
- Crystal, D.** 1986. ‘The ideal dictionary, lexicographer and user’ in R. F. Ilson (ed.), *Lexicography: An emerging international profession: 72–81*. Manchester: MUP.
- De Schryver, G-M.** 1999. Bantu Lexicography and the Concept of *Simultaneous Feedback*, Some preliminary observations on the introduction of a new methodology for the compilation of dictionaries with special reference to a bilingual learner’s dictionary *Cilubà-Dutch*. Unpublished MA thesis. Ghent: Ghent University.
- De Schryver, G-M.** 2003a. Lexicographers’ Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16.2: 143–199.
- De Schryver, G-M.** 2003b. Online Dictionaries on the Internet: An Overview for the African Languages. *Lexikos* 13 (AFRILEX-reeks/series 13: 2003): 1–20.
- De Schryver, G-M and Prinsloo, D. J.** 2000. ‘Dictionary-Making Process with “Simultaneous Feedback” from the Target Users to the Compilers’ in U. Heid et al. (eds.): 197–209.
- De Schryver, G-M and Prinsloo, D. J.** 2001. Fuzzy SF: Towards the ultimate customised dictionary. *Studies in Lexicography* 11.1: 97–111.
- Docherty, V. J.** 2000. ‘Dictionaries on the Internet: an Overview’ in U. Heid et al. (eds.): 67–74.
- Dodd, W. S.** 1989. ‘Lexicomputing and the dictionary of the future’ in G. James (ed.), *Lexicographers and Their Works: 83–93*. Exeter: EUP.
- Dolezal, F. T. and McCreary, D. R.** 1999. *Pedagogical Lexicography Today. A Critical Bibliography on Learners’ Dictionaries with Special Emphasis on Language Learners and Dictionary Users*. Tübingen: Max Niemeyer Verlag.
- ELDIT.** 2002. Elektronisches Lern(er)wörterbuch Deutsch – Italienisch. Available from: <<http://dev.eurac.edu:8081/MakeEldit1/Eldit.html>>.
- Harley, A.** 2000. ‘Cambridge Dictionaries Online’ in U. Heid et al. (eds.): 85–88.

- Heid, U. et al.** (eds.) 2000. *Proceedings of EURALEX 2000*. Stuttgart: IMS, Universität Stuttgart.
- Hulstijn, J. H.** 1993. When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *The Modern Language Journal* 77: 139–147.
- Hulstijn, J. H. and Atkins, B. T. S.** 1998. ‘Empirical research on dictionary use in foreign-language learning: survey and discussion’ in B. T. S. Atkins (ed.): 7–19.
- Hulstijn, J. H. and Trompetter, P.** 1998. ‘Incidental learning of second language vocabulary in computer-assisted reading and writing tasks’ in D. Albrechtsen et al. (eds.), *Perspectives on foreign and second language pedagogy*: 191–200. Odense: OUP.
- Joffe, D. and De Schryver, G-M.** 2004. ‘TshwaneLex – A State-of-the-Art Dictionary Compilation Program’ (see elsewhere in the current Proceedings)
- Knight, S.** 1994. Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal* 78: 285–299.
- Laufer, B.** 2000. ‘Electronic dictionaries and incidental vocabulary acquisition: does technology make a difference?’ in U. Heid et al. (eds.): 849–854.
- Laufer, B. and Hill, M.** 2000. What Lexical Information do L2 Learners Select in a Call Dictionary and How Does it Affect Word Retention? *Language Learning & Technology* 3.2: 58–76.
- Leech, G., Rayson, P. and Wilson, A.** 2001. *Companion Website for: Word Frequencies in Written and Spoken English: based on the British National Corpus*. Available from: <<http://www.comp.lancs.ac.uk/ucrel/bncfreq/>>.
- Löfberg, L.** 2002. Miksi sanat eivät löydy sanakirjasta? Tapaustutkimus MOT Englanti 4.0. Unpublished MA thesis. Tampere: University of Tampere.
- Nesi, H.** 2000. ‘Electronic Dictionaries in Second Language Vocabulary Comprehension and Acquisition: the State of the Art’ in U. Heid et al. (eds.): 839–847.
- Popescu-Belis, A., Armstrong, S. and Robert, G.** 2002. ‘Electronic Dictionaries – from Publisher Data to a Distribution Server: the DicoPro, DicoEast and RERO Projects’ in *Proceedings of LREC 2002*: 1144–1149.
- Prószycki, G. and Kis, B.** 2002. ‘Development of a Context-Sensitive Electronic Dictionary’ in A. Braasch and C. Povlsen (eds.), *Proceedings of EURALEX 2002*: 281–290. Copenhagen: Center for Sprogteknologi, Københavns Universitet.
- Sato, H.** 2000. ‘Multi-Functional Software for Electronic Dictionaries’ in U. Heid et al. (eds.): 863–870.
- Sobkowiak, W.** 1999. *Pronunciation in EFL Machine-Readable Dictionaries*. Poznań: Motivex.
- Warburton, Y.** 2000. The Oxford English Dictionary – From OED to OED Online. *International Journal of Lexicography* 13.2, *EURALEX Newsletter Summer 2000*: 7–8.
- Wiegand, H. E.** 1998. *Wörterbuchforschung: Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie. 1. Teilband*. Berlin: Walter de Gruyter.