HAL
open science

Dictionaries as aids for language learning

Alex Boulton, Sylvie de Cock

To cite this version:

Alex Boulton, Sylvie de Cock. Dictionaries as aids for language learning. Patrick Hanks & Gilles-Maurice de Schryver. International Handbook of Lexis and Lexicography., Springer, pp.XX-XX, 2017. hal-00937985

HAL Id: hal-00937985

https://hal.science/hal-00937985v1

Submitted on 7 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Alex Boulton & Sylvie De Cock. (2017). Dictionaries as aids for language learning. In P. Hanks & G.-M. de Schryver (eds), *International Handbook of Lexis and Lexicography*. New York: Springer, n.p. DOI: 10.1007/978-3-642-45369-4_25-1.

Abstract

Research has consistently shown that a few thousand word families can be enough to understand most written or spoken text (e.g. Nation, 2006; Van Zeeland & Schmitt, 2013). The implication is that these frequent items need to be learned – and to be learned thoroughly. However, diminishing returns set in, and accessible tools are required for the many tens of thousands of other items. Dictionaries are among the most widely-used tools for foreign and second (L2) language learning, and can help with both frequent and infrequent items. This chapter begins with a discussion of dictionaries themselves in relation to L2 use, and the main issues affecting their development over the last few decades in particular. These include the increasing use of empirical data, especially in the form of corpora, and the appearance of monolingual learner dictionaries in addition to bilingual and other dictionary types. It then moves on to a discussion of research into how dictionaries are used by non-native speakers (NNSs), especially as a reference resource for encoding and decoding, but also their impact on language learning itself.

1. Introduction: Historical overview and major issues

As will have become clear in this *Handbook*, dictionaries have a long and varied history. In essence, their intention is to inform people unsure of the meanings or uses of 'words', such people often being non-native speakers (NNSs). From antiquity, word repositories included *ad hoc* collections of translations intended to help with commerce and conquest; from the Middle Ages, the margin glosses produced by monks gave rise to more principled collections for deciphering religious texts in Latin, Greek and other classical languages. The Enlightenment saw the development of dictionaries for native speakers (NSs), thus giving rise to a preoccupation with 'good usage', a trend which was only questioned in the 20th century with a more descriptive approach to lexicography, which has particular implications for learners and other NNSs. The first part of this chapter looks at recent (i.e. 20th and 21st century) developments in the design, compilation and use of various types of dictionaries and associated resources for and by NNSs. The second part moves on to research conducted on dictionary use and how this informs lexicography as well as informing good practice for dictionary use by teachers, learners and other users.

1.1. A descriptive approach to word lists in language learning

The descriptive approach to NNS lexicography predates the modern corpus era. The first notable event was the appearance of Thorndike and Lorge's *Teacher's Wordbook of 30,000 Words* which was published in 1944, though earlier versions had appeared in 1921 and 1931. Based on a collection of 18 million words, the book contained 30,000 entries listed according to frequency bands per million words in different categories, including juvenile literature and magazines. Lorge went on to collaborate in West's *General Service List of English Words*, the definitive version of which was published in 1953 (earlier versions date back to 1936). West's work was based on a much smaller corpus (2.5 million words in the main) and provided far fewer target items, but the volume included not just frequencies but also meanings. The 2,000 entries include lexical 'families', such that the various meanings and forms of *present*

and *presence* as a verb, noun and adjective (i.e. inflections and transparent derivations) are included under the same heading. The overall frequency is given, along with percentage frequencies for each main meaning of each main form, and the list is still available and used today (e.g. at http://www.lextutor.ca/freq/lists_download). While this frequency approach was common in English, substantial work was also conducted in some other languages, notably French with the *Dictionnaire Fondamental de la Langue Française* (Gougenheim, 1958), which marked a turning point by making significant use of spoken language. But it is only recently that computer tools have made the task considerably easier, giving rise for example to the Routledge *Frequency Dictionaries* currently available for 14 major languages. As these span nearly a decade, there are a number of differences between volumes; but the French version, as an example (Lonsdale & Le Bras, 2011), gives the 5,000 most frequent lemmas in rank order derived from a corpus of over 20 million words (evenly divided between speech and writing), along with part of speech, an example sentence from the corpus plus English translation, and thematic selections and alphabetic lists for ease of use, etc. Though the series is called 'dictionaries', it is rather far from the prototypical dictionary many learners are used to – though not as distant as some other frequency lists such as the *Academic Word List* (Coxhead, 2000) which was based on the GSL, the *Academic Formulas List* (Simpson-Vlach & Ellis, 2010), or the *Phrasal Expressions List* (Martinez & Schmitt, 2012), all of which serve to complement dictionaries in aiding the learner to decide what items (and which meanings and uses) to give priority to in different areas.

1.2. Intuition and data in learner dictionaries

Returning to more standard dictionaries, the first major innovation of the 20th century was probably the appearance of Hornby's seminal work which is now in its eighth edition as the *Oxford Advanced Learner's Dictionary*, accompanied by an American edition and various spinoffs for phrasal verbs, idioms, and other language areas. This derived from the insight that dictionaries intended for NSs often were of limited relevance to language learners, an insight that affected choices of what items to include (the most frequent, but also the most useful and with the widest coverage), what information to provide (as regards meaning, form and usage), as well as how to present it (with more comprehensible definitions and suitable examples). The world of monolingual learner's dictionaries (MLDs) in particular, however, was to undergo a fundamental revolution (Hanks, 2012) in the 1980s with the appearance of the *Collins Cobuild English Language Dictionary*, based on the Bank of English corpus as conceived by John Sinclair, which has developed from the original 7 million words to over 650 million today as a subset of the 4.5 billion-word Collins Corpus. The basic idea was that lexicographers should work with systematic collections of language data (corpora) rather than their intuitions or fortuitous examples of salient items, hence its motto "helping learners with real English" (see the papers in Sinclair, 1987, for an overview of the original project). This was essential in deciding what words to include, as well as in ordering entries by the most frequent meanings and uses, but beyond this it notably featured definitions in full sentences without abbreviations, and examples extracted directly from the corpus itself. (See also de Schryver, 2012, for a history of lexicography based on a corpus of academic publications.)

Even today, Hanks (2013, p. 145) can argue that definitions in many dictionaries "owe more to art than to science" as lexicographers seek to come to terms with the fuzzy nature of

meaning and the probabilistic nature of usage. The very fact that dictionaries select different items to which they attribute different meaning classifications presented in different orders is testimony to the fact that entirely objective answers have yet to be achieved; the orderly presentation of dictionaries may even mislead the user into believing that language itself is orderly. Nonetheless, other publishers soon followed the corpus lead with their own MLDs, to the extent that the major ones for English today are all corpus-based to a greater or lesser degree: Cobuild, Oxford, MacMillan, Longman and Cambridge, as well as Merriam-Webster for American English (several of the British publishers also now have American editions, and all include information on major varieties of English). Indeed, attempts to avoid corpora in major lexicographical enterprises today would be "laughed out of court" (McCarthy, 2008, p. 564). Current MLDs generally feature more-or-less authentic examples, information about frequency, help boxes for close synonyms and confusing items, and special entries for common errors or problems, thematic lists, colour pictures and maps, and appendices ranging from verb tables to thesauruses to writing guides. Research involving learner corpora (i.e. corpora that contain written and/or spoken productions by language learners) has also fed into some of the materials included in a number of MLDs, typically in the form of error notes and writing guides. User-friendliness is key as lexicographers are well aware (cf. Nielsen, 2008), and is reflected in limited metalanguage and a core defining vocabulary (first rigorously introduced by Longman), as well as in attempts to signpost longer entries (cf. Lew, 2010). They come in all shapes and sizes, from small pocket dictionaries to desk volumes of 100,000 entries or more. Most include a CD/DVD-ROM providing extra information (e.g. further examples from the source corpus), and have an online version which is usually accessible free (if without some of the extras available in the print purchase), as well as editions for mobile phones and whiteboards for classroom use; pocket dictionaries are still in vogue in some parts of the world (e.g. Asia; Jian *et al*., 2009). These new interfaces have vastly increased the flexibility and mobility of dictionary consultation, though in the case of most major dictionaries the changes are mainly limited to the quantity of information available (especially access to more examples plus some extra features such as study guides), and speed and ease of access to the same contents as in a paper dictionary (see Rizo-Rodríguez, 2008, for a review).

Though the main electronic dictionaries tend to resemble their printed versions closely, individuals or small teams have worked on a number of entirely new formats. Major initiatives tend to be research-driven rather than learner-focused (e.g. WordNet – see Piasecki *et al.*, 2013), but worthy of a special mention in this chapter are 'pattern dictionaries', where recurring syntactic patterns tend to reveal common meanings (Hanks & Pustejovsky, 2005); 'organic dictionaries' derived from self-selecting collocations arising from a specialised corpus (Williams, 2012); and various tools for creating personal or collaborative 'word banks' or wiktionaries which can be suitable for general or LSP classes (Horst & Cobb, 2001). Online dictionaries can also be integrated directly into learning platforms: with Hypertext (http://www.lextutor.ca/hypertext), for example, after submitting a text the user can click on a word to bring up the dictionary entry for that item (Cobb, 2014).

1.3. Monolingual and bilingual dictionaries

The discussion so far has mainly focused on English and on monolingual dictionaries, and not without reason. Hanks (2013, p. 104, p. 355) happily acknowledges that lexicography for languages in other countries may be more advanced in terms of theoretical foundations (e.g. Russia), or phraseology in dictionaries (e.g. Czech), but it is no secret that dictionary publishing is not a purely altruistic venture – it is also a business in search of profits. As dictionary production is highly competitive (with several major dictionaries in the English MLD market where once there was only the OALD), it has been "marked by rapid and constant change, technological advance, innovative and creative development and response to users' needs" (Kirkness 2004, p. 68). An opposing force though is inertia: given the expense in producing a dictionary, publishers also tend to be conservative and to provide customers with slightly improved versions of tools they are already familiar with (Hanks, 2008). The market for dictionaries of English is enormous compared to that for most other languages, and monolingual versions can be sold around the world while bilingual dictionaries have a much more restricted market. Nonetheless, the bilingual dictionary in one form or another is probably the single most common language tool for NNSs, preferred by learners in almost all studies (e.g. Atkins & Varantola, 1997; but cf. Frankenberg-Garcia, 2011), and topping the list of favourite Internet resources for language (Scheffler, 2007; Todd, 2007). Familiarity and staying within the "comfort zone" are no doubt contributing factors (Kaur & Hegelheimer 2005, p. 298), but in some cases bilingual dictionaries may actually be more effective than MLDs (e.g. Lew, 2004).

In bilingual dictionaries, an entry from the target language (L2) into the mother tongue (L1) is mainly intended as an aid in *de*coding meaning, while an L1=>L2 entry is useful for *en*coding – though as such dictionaries are usually intended for learners of either language, the English=>French section of a bilingual dictionary needs to include information relevant for speakers of both languages and hence for both encoding and decoding purposes. A similar problem exists in MLDs, though greater space can be allotted to each individual entry since the focus is often more on encoding, and equivalent information does not need to be repeated in both bilingual directions. This is not a minor issue, as Frankenberg-Garcia (2012, 2014) shows how example sentences which are appropriate for decoding may not be helpful for encoding purposes, and vice versa. It is particularly problematic given the dictionary's focus on individual words rather than contextualised uses, especially in translation where the bilingual dictionary is the staple tool even though it is no easy task to use it given that we do not translate one word at a time (cf. Teubert, 2004). More recently, corpus work has shown that all normal language use relies heavily on 'chunks', and more recently still that many chunks have psycholinguistic reality (e.g. Conklin & Schmitt, 2012). If words really only have 'meaning potentials' which are realised in actual use, how is a learner to distinguish between possible meanings when the dictionary is limited in the sample contexts it offers? The problem is confounded in production: how is a learner to choose between all the options which dictionaries fail to rule out, i.e. to distinguish *normal* uses from creative and exceptional but rare *exploitations* (Hanks, 2013). As Taylor (2012, p. 100) puts it: "A person armed only with the dictionary and the grammar book could well come up with expressions which are fully grammatical and with meanings which can easily be worked out but which happen not to correspond to what speakers of a language would normally say."

1.4. From dictionary to user

So much for dictionaries themselves, but what of the learners? Clearly it is important to improve dictionaries, but it is also worth asking whether we can "improve the users" (Atkins & Varantola, 1997, p. 1). Most dictionaries intended for NNSs, whether monolingual or bilingual, paper or electronic, are very easy to use on first encounter – at least to the extent that the untrained user is able to derive some benefit from them at a very basic level. It may however be precisely *because* of this that training in dictionary optimisation is rare even though it is widely acknowledged to be extremely valuable: "The dictionary is an excellent tool in the hands of a skilled learner. An unskilled user wastes time and comes away frustrated from dictionary consultations" (Roby, 2006, p. 59). Ideally, given the ubiquity of dictionaries of various types, training should be integrated into lessons (Béjoint & Moulin, 1987) early and regularly. Considerable evidence suggests that such training can indeed be effective (e.g. Ranalli, 2013, found a large and significant increase in the number of look-ups following dictionary instruction), but rather than treating it as an add-on activity separate from regular work, Frankenberg-Garcia (2011) and others propose dictionary training to help learners on an *ad hoc* basis whenever the need arises. In this view, we begin not with the dictionary and its hypothetical uses but with the learners themselves and the questions they genuinely have.

Teacher and learner representations of vocabulary in general and dictionary use in particular are not always well founded (cf. Folse, 2004). One prevalent assumption is that dictionaries themselves contain all and only 'right' answers (hence the comment lament: "But I found it in the dictionary"; cf. the 'blind faith' learners have in dictionaries reported by Bowker, 1999), such that any error following dictionary consultation is reputedly the learner's fault. Clearly though, lexicographers have to make choices about what items to include, condensing a huge wealth of data into short entries on the most common features which may therefore not contain the answer to a specific question – especially for phraseology and collocation, or genre-specific uses (Hanks, 2013). Another common assumption is that dictionaries are easy for any literate person to use, such that any specific training is seen as a waste of classroom time and almost an insult to the learner's intelligence – again putting the onus on the learner. A third is that dictionary consultation leads to learning: though there is some evidence for this (e.g. Hulstijn *et al*., 1996; Ronald, 2002), there is no fundamental reason to suppose that a dictionary look-up will necessarily lead to retention, or that an excellent reference tool will necessarily make the best learning aid. At the same time, it seems clear that learners generally look only for 'unknown' words and ignore frequent items which they in fact know only partially and for which dictionaries could be of help; that they tend to seek to add to the number of words they know (overwhelmingly nouns) rather than increasing depth of knowledge of the smaller number of common items which have much more complex meaning potentials and uses; that they stop at the first part of the entry or the first familiar item and do not explore longer entries (the so-called 'kidrule'); that they do not exploit phraseological information (such as it is) in the dictionary; that they find it difficult to understand and thus tend to ignore guides to pronunciation, syntax, register, and so on – let alone the introduction and other supplementary material (see Nesi & Haill, 2002). But all of these representations can be subject to research of one type or another, which is the topic of the next section.

2. Description of research to date

Because dictionary making is a commercial enterprise, "research conducted by dictionary publishers is not generally made public" (Nesi, 2014, p. 39). There is however an increasing quantity of empirical studies on dictionary use by learners. These are highly diverse and target various aspects of dictionary use such as learners' beliefs and attitudes, their look-up behaviour and preferences (e.g. menus vs signposts in long entries for polysemous items), how successful they are with dictionaries as a reference resource (for decoding and/or encoding), and whether dictionary use does indeed translate into learning. Investigating dictionary use is a complex and challenging undertaking in view of the many variables involved (see Nesi, 2014, for a historical overview). Most studies are therefore carried out in experimental settings which are often highly controlled, some even making use of purpose-built dictionary entries or eye-tracking technology, for example.

Because of the complexity involved, findings from different investigations are seldom directly comparable and often raise more questions than they answer (Tono, 2001; Abecassis, 2007; Chan, 2012). The contribution of dictionaries is often analysed by comparing learners' performance on a series of tasks with or without access to dictionaries or dictionary entries; contrasts are sometimes extended to other reference tools, such as comparing performance using dictionaries or concordances derived from computerised corpora (e.g. Boulton, 2010). It is notable that there is as yet no meta-analysis of learners' use of dictionaries *per se*, neither as a learning aid nor as a reference resource for encoding or decoding (the nearest thing to date seems to be Abraham, 2008, with a meta-analysis of computer-mediated glosses). From a synthetic reading though, what clearly emerges from the research is that many aspects of dictionary use are influenced by learner variables such as proficiency level in the L2, whether or not the L2 is also the medium of instruction, the degree scheme (e.g. language majors or students needing languages for specific or academic purposes), familiarity with certain types of dictionaries, cultural and educational environment, and so on (Abecassis, 2007; Laufer & Hill, 2000; Nesi, 2014; Tono, 2001). The focus of the rest of this section is on insights from studies that explore the use of dictionaries as a resource for decoding and/or encoding, and the role they may have in learning vocabulary or other aspects of the L2.

2.1. Dictionaries for decoding

One of the key roles of dictionaries as a reference tool is to help learners understand the meaning of words or expressions in decoding activities (e.g. reading a text in the target language). On the whole, research suggests that learners tend to be rather reluctant to turn to dictionaries in receptive activities (Atkins & Varantola, 1997; Hulstijn *et al*., 1996; Lew, 2004). This finding has a very practical impact on studies that specifically set out to analyse the effectiveness of (various types of) dictionaries in receptive use as they presuppose actual dictionary use. As a result, researchers have had to resort to experimental set-ups that require the participants to use dictionaries (e.g. the use of pseudo-words and made-up entries and/or careful monitoring: Tono, 2001; Lew, 2004). After reviewing findings from empirical studies that compare learners' vocabulary comprehension with and without dictionaries, Chen (2011b, p. 218) concludes that there is "no simple yes-or-no answer to the question whether the dictionary is useful for vocabulary comprehension or not". That said, studies that rely on tests that target actual vocabulary comprehension rather than general

text comprehension (where not understanding some vocabulary items might not be a problem; see also Nesi & Meara, 1991) point to the positive influence of dictionary consultation. The effectiveness of dictionary use has been connected with individual learners' proficiency levels and dictionary reference skills, the relevance, complexity and quality of the information provided, the type of dictionary consulted and the accessibility of information in (long) entries (Tono, 2001; Lew, 2004; Chen, 2011b). Another factor is the type of dictionary used. For example, Lew's (2004) study suggests that bilingual dictionaries are both favoured by the learners and more effective than MLDs as they provide an immediate solution in the L1. It is worth noting though that this preference may arguably have been induced – at least in part – by the use of translation tasks in the experimental set-up. Chen (2010) and Dziemianko (2010) have shown that the electronic medium can speed up and boost the reference support learners get from dictionaries in vocabulary comprehension.

Two examples will exemplify dictionary studies of this type. Chan (2012) investigates advanced ESL learners' use of MLDs to determine the meaning of familiar polysemous words used in less familiar contexts. The participants were asked to select the most appropriate meaning of italicised strings of words in sentences from a list of five choices, first without and then with access to an MLD (the target lexical items for consultation were underlined, e.g. *The minister is alive to the responsibility of his position*). Overall, the results of the multiple-choice tests reveal an improvement in test results after dictionary consultation. The study incorporated self-reporting protocols to uncover learners' strategies and problems when extracting information from the MLDs. For example, the participants frequently reported that there were "too many definitions and examples in the entries" and that "the arrangement of the entries was too long and confusing" (Chan, 2012, p. 128). Chan also found that learners tend to be rather insensitive to grammatical information in the entries when decoding, which sometimes hindered the comprehension of the familiar words in unfamiliar grammatical contexts.

The definition component of entries in MLDs is often presented as the key element in vocabulary comprehension (Laufer, 1993). To test this, Frankenberg-Garcia (2012) set out to compare Portuguese EFL learners' comprehension of low-frequency words (1) without any reference consultation, (2) with the help of a dictionary definition only, or with (3) a single or (4) multiple well-chosen corpus examples. The results of the multiple-choice test (which required the participants to select the best Portuguese translation of the items under study) show that the performance of the learners who used multiple examples was very similar to that of the learners who consulted the definition. A quasi-replication study (Frankenberg-Garcia, 2014) with secondary-school learners came to similar conclusions, additionally showing that the *types* of examples needed for encoding and decoding purposes are very different, but that multiple examples can in fact be of greater help than dictionary definitions when it comes to language production. Though these examples were manually selected, it should be possible to establish criteria for appropriate examples (cf. Hanks, 2013, p. 410), and to implement these at least partly automatically; steps have been taken in this direction by the SketchEngine team for Macmillan (see Rundell & Kilgarriff, 2011).

This is a pre-publication version. For the version of record, please see DOI: 10.1007/978-3-642-45369-4_25-1
(or email me at alex.boulton@atilf.fr)

2.2. Dictionaries for encoding

Although MLDs have often been presented as encoding tools *par excellence* by lexicographers and researchers (Rundell, 1999; McEnery *et al*., 2006; De Cock & Granger, 2004), learners' actual use of these dictionaries for production can appear to be rather limited or unsuccessful (Lew, 2004). Frequently cited explanations for this include learners' lack of awareness of their own needs and problems when encoding (e.g. Frankenberg-Garcia, 2011), including and perhaps especially in terms of collocation (e.g. Laufer, 2010). They similarly lack awareness of the wealth of useful encoding information in this type of dictionary and have poor dictionary skills, both connected with a general lack of dictionary training. And of course, sometimes the dictionaries themselves may be responsible if they do not include the necessary words or information, or if the answers are difficult to find or buried among a mass of information which is poorly signposted (e.g. Chan 2012).

One of the typical features of corpus-informed MLDs is the inclusion in entries of a series of example sentences or phrases taken from computerised corpora, either directly or subject to editing in some cases. These examples, which illustrate the use of words or strings of words in different contexts, appear to be particularly useful to learners in production activities provided the focus is on examples specifically meant to help production (Frankenberg-Garcia, 2012, 2014), and favour a usage-based model of language learning (Taylor, 2012). Research also suggests that learners might benefit from direct or mediated access to computerised corpora and other online resources when producing texts (although there appear to be strong individual differences; e.g. Charles, 2012), as this provides them with access to the use of target words in more, different and even highly specialised contexts which might be close to the contexts in which they need to use the words. As a result, the combination of an electronic dictionary with other electronic resources, typically corpora or concordances, has been widely advocated and occasionally implemented (Paquot, 2012; Verlinde & Peeters, 2012). Some degree of corpus mediation may be required, since learners (especially at lower levels of proficiency and with limited needs and ambitions) cannot be expected to be corpus linguists any more than they can be expected to be lexicographers (Chambers, 2005, p. 114). Further, direct corpus consultation can be a challenge even for experienced corpus users when they need immediate answers to production questions (Frankenberg-Garcia, 2012; Chambers & Kelly, 2004).

As mentioned above, dictionary use and learner variables are closely intertwined, and the role of learners' L1 in L2 production cannot be ignored (Bogaards, 2010). Learners can often be seen to access the L2 via their L1 when encoding, and this is arguably reflected in their reported general preference for bilingual dictionaries. Using bilingual dictionaries for encoding is generally not regarded as desirable, however, insofar as bilingual dictionaries tend to lack the rich production-oriented lexico-grammatical information typically included in MLDs. Bilingualised dictionaries (cf. Bogaards & Hannay, 2004) attempt a solution to this as they seek to "combine the best of both worlds" by bringing together "the extensive knowledge that is condensed in modern monolingual learner's dictionaries and the exploitation of bilingual equivalence that are so well established in the learner's mental lexicon" (Bogaards, 2010, p. 119).

Laufer and Levitzky-Aviad (2006) compare the 'usefulness' of four types of dictionaries in a productive Hebrew into English translation task performed by Hebrew EFL learners: (1) a printed bilingualised (L1-L1-L2) dictionary; (2) a printed L1-L2 bilingual dictionary; (3) a printed L1-L2-L2 bilingualised dictionary (referred to as bilingual dictionary plus or BD+); (4) a computerised BD+. The translations were scored taking both semantic and syntactic appropriateness into consideration. The results suggest that the paper and electronic BD+ can have advantages over other dictionaries. It would be interesting to see if this advantage is confirmed in future studies focusing on encoding tasks that do not directly involve translation (e.g. writing).

Although bilingualised dictionaries appear to be commonly used in some countries (e.g. China: Chen, 2011a), they are still generally less easily accessible than traditional bilingual dictionaries and MLDs. An alternative is the production dictionary (e.g. the *Longman Language Activator*) which is specifically designed to help with writing. More recently, the *Louvain English for Academic Purposes Dictionary* is presented as a "dictionary-cum-writing aid" (Paquot, 2012, p. 163), and makes full use of the flexibility of the electronic medium to admit considerable customisation. French-speaking users have access to L2 entries via their L1 and language use needs are taken into consideration (e.g. for EAP learners and in disciplines such as medicine, economics, linguistics or political sciences). Within the framework of the project, an examination of the log files is currently under way to determine how the tool could be improved to better meet the users' needs.

2.3 Dictionaries and language learning

As pointed out by Lew and Doroszewska (2009, p. 239), learning new vocabulary can be regarded as "a useful by-product of dictionary consultation". Any learning resulting from the use of dictionaries as reference tools has been labelled as incidental learning (as opposed to intentional learning: Hulstijn, 2003; Bogaards, 2010). In empirical studies of vocabulary learning and dictionary use, learning tends to be operationalised as the extent to which new words, meanings or collocations are remembered by the learners in immediate and/or delayed (unexpected) vocabulary tests following decoding or encoding tasks involving the use of a dictionary. Besides comparing learners' retention of words and collocations with and without dictionary use, various studies also explore the impact of a series of dictionary-related factors on retention: types of dictionary used, dictionary medium (paper vs electronic), use of various entry components (e.g. L1 definitions, L2 translations, examples of use, animated pictures, etc.), look-up behaviour or number of look-ups. On the whole, dictionary consultation seems to have a positive impact on the retention of the items concerned (e.g. Laufer & Hulstijn, 2001). This impact appears to be enhanced by factors such as the presence of word-focused activities and the highlighting of the target items in the tasks carried out by the learners, or the availability of an L1 translation in the dictionary entries (Hill & Laufer, 2003; Lew & Doroszewska, 2009; Chen, 2011b).

Hill and Laufer (2003) investigated the effect of task type on incidental L2 vocabulary learning. The Chinese ESL learners in the study had to (1) read and understand a text (containing unfamiliar target words, which were highlighted in the text and whose meanings could be looked up in a dictionary); (2) perform one of three tasks (a form-oriented production task – selecting one of four target words that means the same as a given

synonym or paraphrase of the target word; a form-oriented comprehension task – selecting the meaning of a given target word from a list of four options; or a meaning-oriented task – answering yes/no text comprehension questions); and then (3) take immediate and delayed vocabulary tests (they had to provide the meaning of the target words in English or Chinese). The results reveal that the two form-oriented tasks (and especially the productive task) were more conducive to vocabulary learning (i.e. better retention) than the meaning-oriented task because they were more demanding for the learners and led to more look-ups. Hill and Laufer conclude that "an important factor determining task effectiveness for vocabulary learning is the amount of word-related activity that the task induces" (p. 87).

One recurrent theme in research into dictionary use and vocabulary learning is the connection between depth of processing and retention. Findings seem to support the idea that the deeper the processing, the better the retention of new vocabulary (cf. the Involvement Load Hypothesis: Laufer & Hulstijn, 2001). However, as pointed out by Dziemianko (2012) in a discussion of the effects of paper vs electronic dictionaries on word retention, deeper processing should not be equated with laborious look-up processes typically associated with printed dictionaries, as retention can be higher when using electronic rather than paper-based dictionaries. Dziemianko argues that depth of processing is triggered by attention, which "does not necessarily correlate with the effort put into dictionary look-up, which is often performed quite automatically. Instead, the conspicuousness of headwords and entries on the computer screen might arrest users' attention" (p. 331). This is arguably in line with Schmidt's noticing hypothesis (1990). Although he and others in the area have abandoned the original 'strong' form of the hypothesis (whereby "noticing is the necessary and sufficient condition for converting input to intake"; p. 129), it does seem that conscious awareness of new language can promote learning. And since dictionary use involves conscious effort, it may indeed contribute to language learning.

As we saw above, the electronic versions of some MLDs (e.g. Longman, Oxford, Macmillan) have been transformed into veritable language learning environments featuring a whole host of resources, exercises, (learner-corpus informed) error notes, and grammar and writing sections. To our knowledge, however, no empirical research has been published into the use of this extra material by learners (or teachers) and its impact on actual language learning. The web-based Interactive Language Toolbox (Verlinde & Peeters, 2012) is an example of what Abel (2010) calls dictionary-cum-CALL systems, which combine a dictionary and a learning tool (for French and Dutch vocabulary in this case). Verlinde and Peeters (2012) explain how insights from a usability study (based on a think-aloud experiment and a users' questionnaire) were taken into consideration to enhance and simplify data access and search facilities. They argue that "users' expectations and habits are still insufficiently taken into account by lexicographers" and highlight the importance of implementing a "more user-oriented lexicography" (Verlinde & Peeters, 2012, p. 161–162).

Conclusion

We have seen in this chapter that dictionaries are a perennial staple for language learners and foreign or second language users of all types. Bilingual dictionaries remain the basis for this, though curiously they have been less innovative perhaps than MLDs. Though there may

be a move from print to electronic formats, this reliance on dictionaries seems unlikely to disappear any time soon as they also top the list of students' preferred language technologies (Steel & Levy, 2013). However, they are inevitably complemented by other resources which are likely to become even more widely accessed as users become more internet-literate and resort to on-line translators and corpora, or indeed general-purpose search engines which enable them to consult the 'web-as-corpus' (see Boulton, 2015).

Though research into dictionary use is a surprisingly recent phenomenon (Nesi, 2000), the results seem to point consistently in the same direction: dictionaries (if appropriately chosen and appropriately used) can contribute substantially to enhancing language performance in terms of both encoding and decoding, and can lead to retention of new items (meaning, form and usage). A key term however is 'appropriate' – dictionaries tend to be taken for granted in many instructional settings, and training in their efficient use is arguably one of the unaccountably missing ingredients in language teaching today.

It would be an unusual paper in applied linguistics which did not call for more research, but this is perhaps particularly true in the area of dictionary use by L2 users. Target populations need to be extended beyond the favoured audience of university students, and include younger secondary-school learners as well as adult users who need an L2 on a more or less regular basis in their professional or social lives. There is also a particular lack of ecological studies on actual dictionary use over an extended period (as opposed to controlled, experimental or laboratory designs where the participants perform an imposed task, often in just a few minutes). Such studies may be more difficult to organise rigorously given the dynamic and complex nature of language learning (cf. Verspoor *et al*., 2011), but should become increasingly accessible as more learners turn to electronic resources which enable their look-ups to be unobtrusively tracked systematically at distance and evaluated in terms of outcomes in language use.

References

Abecassis, M. (2007). Is lexicography making progress? On dictionary use and language learners' needs. *Lexikos (AFRILEX-series), 17*, 247–258.

Abel, A. (2010). Towards a systematic classification framework for dictionaries and CALL. In S. Granger & M. Paquot (eds.), *eLexicography in the 21st century: New challenges, new applications*. Louvain-la-Neuve: Presses universitaires de Louvain, pp. 3–12.

Abraham, L. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning, 21*(3), 199–226.

Atkins, S. & Varantola, K. (1997). Monitoring dictionary use. *International Journal of Lexicography, 10*(1), 1–45.

Béjoint, H. & Moulin, A. (1987). The place of the dictionary in an EFL programme. In A. Cowie (eds.), *The dictionary and the language learner*. Tübingen: Max Niemeyer, pp. 381–392.

Bogaards, P. (2010). Dictionaries and second language acquisition. In A. Dykstra & T. Schoonheim (eds.), *Proceedings of the 14th EURALEX international congress*. Leeuwarden: Fryske Akademy, pp. 99–123.

Bogaards, P. & Hannay, M. (2004). Towards a new type of bilingual dictionary. In G. Williams & S. Vessier (eds.), *Proceedings of the 11th EURALEX international congress*. Lorient: Université de Bretagne-Sud, pp. 463–474.

Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning, 60*(3), 534–572.

Boulton, A. (2015). Applying data-driven learning to the web. In A. Leńko-Szymańska & A. Boulton (eds.), *Multiple affordances of language corpora for data-driven learning*. Amsterdam: John Benjamins, pp. 267–295.

Bowker, L. (1999). Exploring the potential of corpora for raising language awareness in student translators. *Language Awareness, 8*(3-4), 160–173.

Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning & Technology, 9*(2), 111–125. http://llt.msu.edu/vol9num2/chambers/default.html

Chambers, A. & Kelly, V. (2004). Corpora and concordancing: Changing the paradigm in language learning and teaching? In A. Chambers, J. Conacher & J. Littlemore (eds.), *ICT and language learning: Integrating pedagogy and practice*. Birmingham: University of Birmingham Press, pp. 183–201.

Chan, A. (2012). The use of a monolingual dictionary for meaning determination by advanced Cantonese ESL learners in Hong Kong. *Applied Linguistics*, *33*(2), 115–140.

Charles, M. (2012). 'Proper vocabulary and juicy collocations': EAP students evaluate do-it-yourself corpus-building. *English for Specific Purposes, 31*, 93–102.

Chen, C. (2011a). Studies on bilingualized dictionaries: The user perspective. *International Journal of Lexicography*, *24*(2), 161–197.

Chen, C. (2011b). Dictionary use and vocabulary learning in the context of reading. *International Journal of Lexicography, 25*(2), 216–247.

Chen, Y. (2010). Dictionary use and EFL learning: A contrastive study of pocket electronic dictionaries and paper dictionaries. *International Journal of Lexicography*, *23*(3), 275–306.

Cobb, T. (2014). A resource wish-list for data-driven learning in French. In H. Tyne, V. André, A. Boulton, C. Benzitoun & Y. Greub (eds.), *Ecological and data-driven perspectives in French language studies*. Cambridge: Cambridge Scholars, pp. 257–292.

Conklin, K. & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics, 32*, 45–61.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238.

De Cock, S. & Granger, S. (2004). Computer learner corpora and monolingual learners' dictionaries: The perfect match. *Lexicographica, 20*, 2–16.

de Schryver, G-M. (2012). Trends in twenty-five years of academic lexicography. *International Journal of Lexicography, 25*(4), 464–506.

Dziemianko, A. (2010). Paper or electronic? The role of dictionary form in language reception, production and the retention of meaning and collocations. *International Journal of Lexicography*, *23*(3), 257–273.

Dziemianko, A. (2012). On the use(fulness) of paper and electronic dictionaries. In S. Granger & M. Paquot (eds.), *Electronic lexicography*. Oxford: Oxford University Press, pp. 319–341.

Folse, K. (2004). *Vocabulary myths: Applying second language research to classroom teaching*. Ann Arbor: University of Michigan Press.

Frankenberg-Garcia, A. (2011). Beyond L1-L2 equivalents: Where do users of English as a foreign language turn for help? *International Journal of Lexicography, 24*(1), 97–123.

Frankenberg-Garcia, A. (2012). Learners' use of corpus examples. *International Journal of Lexicography, 25*(3), 273–296.

Frankenberg-Garcia, A. (2014). The use of corpus examples for language comprehension and production. *ReCALL, 26*(2) : 128–146.

Gougenheim, G. (1958). *Dictionnaire fondamental de la langue française*. Paris: Didier.

Hanks, P. (2008). The lexicographical legacy of John Sinclair. *International Journal of Lexicography, 21*(3), 219–229.

Hanks, P. 2012. The corpus revolution in lexicography. *International Journal of Lexicography, 25*(4), 398–436.

Hanks, P. (2013). *Lexical analysis: Norms and exploitations*. Cambridge, MA: MIT Press.

Hanks, P. & Pustejovsky, J. (2005). A pattern dictionary for natural language processing. *Revue Française de Linguistique Appliquée, 10*(2), 63–82.

Hill, M. & Laufer, B. (2003). Type of task, time-on-task and electronic dictionaries in incidental vocabulary acquisition. *International Review of Applied Linguistics*, *41*, 87–106.

Horst, M. & Cobb, T. (2001). Growing academic vocabulary with a collaborative on-line database. In B. Morrison, D. Gardner, K. Keobke & M. Spratt (eds.), *ELT perspectives on IT & multimedia*. Hong Kong: Hong Kong Polytechnic University, pp. 189–225.

Hulstijn, J. (2003). Incidental and intentional learning. In C. Doughty & M. Long (eds.), *The handbook of second language acquisition*. Oxford: Blackwell, pp. 349–381.

Hulstijn, J., Hollander, M. & Greidanus, M. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *Modern Language Journal, 80*, 327–339.

Jian, J-Y., Sandnes, F., Law, K., Huang, Y-P. & Huang, Y-M. (2009). The role of electronic pocket dictionaries as an English learning tool among Chinese students. *Journal of Computer-Assisted Language Learning, 25*, 503–514.

Kaur, J. & Hegelheimer, V. (2005). ESL students' use of concordance in the transfer of academic word knowledge: An exploratory study. *Computer Assisted Language Learning, 18*(4), 287–310.

Kirkness, A. (2004). Lexicography. In A. Davies & C. Elder (eds.), *The handbook of applied linguistics*. Oxford: Blackwell.

Laufer, B. (1993). The effect of dictionary definitions and examples on the use and comprehension of new L2 words. *Cahiers de Lexicologie, 63*(2), 131–142.

Laufer, B. (2010). The contribution of dictionary use to the production and retention of collocations in a second language. *International Journal of Lexicography*, *24*(1), 29–49.

Laufer, B. & 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.

Laufer, B. & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*(1), 1–26.

Laufer, B. & Levitzky-Aviad, T. (2006). Examining the effectiveness of 'Bilingual Dictionary Plus': A dictionary for production in a foreign language. *International Journal of Lexicography*, *19*(2), 135–155.

Lew, R. (2004). *Which dictionary for whom? Receptive use of bilingual, monolingual and semi-bilingual dictionaries by Polish learners of English*. Poznan: Motivex.

Lew, R. (2010). Users take shortcuts: Navigating dictionary entries. In A. Dykstra & T. Schoonheim (eds.), *Proceedings of the 14th EURALEX international congress*. Leeuwarden/Ljouwert: Fryske Akademy, pp. 1121–1132.

Lew, R. & Doroszewska, J. (2009). Electronic dictionary entries with animated pictures: Lookup preferences and word retention. *International Journal of Lexicography*, 22(3), 239–257.

Lonsdale, D. & Le Bras, Y. (2011). *A frequency dictionary of French: Core vocabulary for learners*. London: Routledge.

Martinez, R. & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics, 33*(3), 299–320.

McCarthy, M. (2008). Accessing and interpreting corpus information in the teacher education context. *Language Teaching, 41*(4), 563–574.

McEnery, T., Xiao, R. & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London: Routledge.

Nation, I.S.P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review, 63*(1), 59–82.

Nesi, H. (2000). *The use and abuse of EFL Dictionaries: How learners of English as a foreign language read and interpret dictionary entries*. Tübingen: Max Niemeyer.

Nesi, H. (2014). Dictionary use by English language learners. *Language Teaching, 47*(1), 38–55.

Nesi, H. & Haill, R. (2002). A study of dictionary use by international students at a British university. *International Journal of Lexicography, 15*(4), 277–306.

Nesi, H. & Meara, P. (1991). How using dictionaries affects performance in multiple-choice EFL tests. *Reading in a Foreign Language, 8*(1), 631–643.

Nielsen, S. (2008). The effect of lexicographical information cost on dictionary making and use. *Lexicos, 18*, 170–181.

Paquot, M. (2012). The LEAD dictionary-cum-writing aid: An integrated dictionary and corpus tool. In S. Granger & M. Paquot (eds.), *Electronic lexicography*. Oxford: Oxford University Press, pp. 163–185.

Piasecki, M., Szpakowicz, S., Fellbaum, C. & Sandford Pedersen, B. (eds.) (2013). *Wordnets and relations. Language Resources and Evaluation, 47*(3).

Ranalli, J. 2013. Online strategy instruction for integrating dictionary skills and language awareness. *Language Learning & Technology*, *17*(2), 75–99.

Rizo-Rodríguez, A. (2008). Review of five English learners' dictionaries on CD-ROM. *Language Learning & Technology, 12*(1), 23–42.

Roby, W. (2006). The internet, autonomy, and lexicography: A convergence? *Mélanges CRAPEL, 28*, 47–66. http://www.atilf.fr/spip.php?article3692

Ronald, J. (2002). L2 lexical growth through extensive reading and dictionary use: A case study. In A. Braasch & C. Povlsen (eds.), *Proceedings of the 10th EURALEX international congress* [Volume 2]. Copenhagen: Center for Sprogteknologi, Copenhagen University, pp. 765–771.

Rundell, M. (1999). Dictionary use in production. *International Journal of Lexicography*, *12*(1), 35–53.

Rundell, M. & Kilgarriff, A. (2011). Automating the creation of dictionaries: Where will it all end? In F. Meunier, S. De Cock, G. Gilquin & M. Paquot (eds.), *A taste for corpora: A tribute to Professor Sylviane Granger*. Amsterdam: John Benjamins, pp. 257–281.

Scheffler, P. (2007). When intuition fails us: The world wide web as a corpus. *Glottodidactica, 33*, 137–145.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158.

Simpson-Vlach, R. & Ellis, N. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics, 31*(4), 487–512.

Sinclair, J. (ed.) (1987). *Looking up: An account of the COBUILD project in lexical computing*. London: Collins.

Steel, C. & Levy, M. (2013). Language students and their technologies: Charting the evolution 2006–2011. *ReCALL, 25*(3), pp. 306–320.

Taylor, J. (2012). *The mental corpus: How language is represented in the mind*. Oxford: Oxford University Press.

Teubert, W. (2004). Units of meaning, parallel corpora, and their implications for language teaching. In U. Connor & T. Upton (eds.), *Applied corpus linguistics: A multidimensional perspective*. Amsterdam: Rodopi, pp. 171–189.

Thorndike, E. & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Columbia University.

Todd, R. (2007). Computer-assisted language use: An internet survey. *CALL-EJ Online, 9*(1), n.p. http://www.tell.is.ritsumei.ac.jp/callejonline/journal/9-1/todd.html

Tono, Y. (2001). *Research on dictionary use in the context of foreign language learning: Focus on reading comprehension*. Tübingen: Max Niemeyer.

Van Zeeland, H. & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics, 34*(4), 457–479.

Verlinde, S. & Peeters, G. (2012). Data access revisited: The Interactive Language Toolbox. In S. Granger & M. Paquot (eds.), *Electronic lexicography*. Oxford: Oxford University Press, pp. 147–162.

Verspoor, M., de Bot, K. & Lowie, W. (eds.) (2011). *A dynamic approach to second language development: Methods and techniques*. Amsterdam: John Benjamins.

West, M. (1953). *A general service list of English words*. London: Longman.

Williams, G. (2012). Bringing data and dictionary together: Real science in real dictionaries. In A. Boulton, S. Carter-Thomas & E. Rowley-Jolivet (eds.), *Corpus-informed research and learning in ESP: Issues and applications*. Amsterdam: John Benjamins, pp. 217–238.

This is a pre-publication version. For the version of record, please see DOI: 10.1007/978-3-642-45369-4_25-1
(or email me at alex.boulton@atilf.fr)