New ways of indicating meaning in electronic dictionaries: hope or hype?

Robert Lew
Adam Mickiewicz University, Poznań, Poland

School of English, Adam Mickiewicz University
Niepodległości 4
PL-61-874 Poznań
Poland

tel. +48 61 8201048
email: rlew@amu.edu.pl

Abstract:

User studies have shown repeatedly that the type of information that dictionary users want from dictionaries the most is meaning. This is not surprising: in fact the dictionary has always been perceived as a repository of meanings of words. While this popular view has not changed much, the advent of electronic dictionaries can change the lexicographers’ approach to the indication of meaning. The traditional way in paper dictionaries has been to explain words with words, using either a definition (typically in the language of the headword) or an equivalent (typically in another language). Only occasionally have pictorial illustrations been used. In contrast, electronic dictionaries open up new possibilities in this regard. Pictorial illustrations, including colour ones (on colour displays) can be included more easily than before; but electronic dictionaries can also offer media objects that paper cannot carry, such as animation or videos. Will the dictionary users, however, always appreciate and benefit from these new ways of indicating meaning? I will examine the existing evidence, and try to offer informed guesses where evidence is not yet available.

1. Meaning: the thing most wanted by users

Different dictionary users have different needs when consulting dictionaries, and those needs – apart from personal preferences – also depend on the specific task in which dictionary users happen to be involved and circumstances of consultation. So much is obvious from past user research, but what we also learn from such research is that – other things being equal – there is one consultation need that clearly dominates across studies of dictionary users, quite independent of the users’ mother tongue, nationality, and proficiency levels. The thing most wanted from dictionaries is meaning. Even dictionary definitions themselves reflect this fact, as seen, for example, in the following definition given under dictionary in the online version of the Cambridge Dictionary for Advanced Learners (CALD2):

\[
\text{dictionary} \\
\text{noun [C]} \\
1\text{ a book that contains a list of words in alphabetical order with their meanings explained or written in another language, or a similar product for use on a computer}
\]

Now, since meaning appears to be such a fundamental element in a lexicographic work, it is critical that meaning is represented in a way that satisfies this predominant need of
the dictionary users. To this end, lexicographers and dictionary editors have traditionally employed a repertoire of devices for representing meaning in paper dictionaries. With the growing role of electronic dictionaries, there is potential for this repertoire to be extended, and also for some of the traditional modes of meaning provision to be used in somewhat different ways. This is the focus of the present paper. I will continue to discuss the provision of meaning in section 3, but first (section 2) let us have a brief look at issues of space in electronic dictionaries, as they have immediate consequences for the delivery of information about meaning.

2. Issues of space in electronic dictionaries

2.1. The two types of space in electronic dictionaries
One statement that seems to be repeated over and over again by many authors discussing electronic dictionaries is that space constraints are no longer a concern in electronic dictionaries. In a forthcoming paper (Lew in press; see also Corréard 2002), I question this generalization as simplistic, arguing that it is necessary to distinguish at least two senses of lexicographic space: storage space, which refers to the totality of lexicographic data that a dictionary can hold; and presentation space, which refers to the set of data that can be displayed to the user at a single time. Space constraints may no longer be much of a concern for the former sense of space, but are certainly valid for presentation space. Restrictions of presentation space become especially acute with the smaller displays. Thus, portable handheld devices, so very popular throughout Asia, have very limited presentation space, much more so than traditional paper pocket dictionaries (especially if you factor in the smaller effective resolution of the typical LCD displays compared with printed paper), and this introduces serious restrictions on how much data the user can view, even if storing massive amounts of data is no longer a problem.

2.2. Storage space of portable dictionaries is still limited and has to be used wisely
The storage capacity of modern electronic devices keeps growing in accordance with Moore’s law, roughly doubling every two years, but there are fundamental limits to every technology, including how densely you can package electronic devices and how small they can be made. At this time, storage is still a non-trivial issue in the smallest electronic dictionaries, including specialized handheld devices, but also lexicographic applications on mobile phones, which are foreseen by some (Piotrowski 2007) to supplant dedicated handheld electronic dictionaries.¹ Quite apart from space concerns, producing meaningful, quality lexicographic content is as costly as ever, still requiring time-consuming professional human involvement. It would not make very good sense to invest lexicographer work time, money, and storage space in something that users do not necessarily appreciate or are not helped by. There remains of course the question of product marketing, where perfectly useless features can still perhaps be used as effective selling points — but recommending such features for inclusion in dictionaries would not be compatible with the user-centred view of lexicography.

2.3. Hyperlinking, exploding entries: improved access to content
One recipe for a happy marriage of the increased storage of modern electronic media with the restricted size and resolution of visual displays is to make the most of the potential of electronic displays for dynamic presentation. Thus, in Lew (in press) I make a case for the more universal use of such devices as entry menus, exploding entries, pop-up windows

¹ This scenario is quite likely; for example, quality English learners’ dictionaries (LDOCE4, CALD2) and a variety of other monolingual and bilingual dictionaries are already available in versions for mobile phones running the Symbian operating system or as Java applications
activated on mouse-hover events – what I call immediate cross-references. A good example of this is the following entry (Figure 1) from LDOCE4 online. When the user wants to access information on the word *damage* by double-clicking on it, the cross-referenced entry pops up on the screen, but the original (source) entry is still retained and largely visible in the background window.

![Figure 1](https://example.com/disaster.png)

**Figure 1.** An example of immediate cross-reference. Longman Dictionary of Contemporary English 4 online entry for *disaster*. The target entry *damage* pops up but the source entry is still retained and partially visible in the background (after Lew in press).

Using entry menus and immediate cross-referencing, we can avoid screen clutter while at the same time giving access to rich content – provided the dictionary has it.

---

3. Representation of meaning in dictionaries

3.1. Verbal representation

Dictionaries are books about words, but they are also books with words: words is what dictionaries explain, but words are also the tool with which to do the explanation. In this sense, the dominant way of indicating meaning in dictionaries has been verbal explanation. Such explanation, however, can take a number of forms which can be used alone or in combination.

3.1.1. Definition

The definition is the most common and prototypical conveyor of meaning in monolingual dictionaries; normally, it consists of a paraphrase of a lexical unit through a more elaborate syntactic construction, although more adventurous variations are not impossible (Hanks 1987; Wierzbicka 1985; Lew and Dziemianko 2006). The definition has survived the test of time, having endured for many centuries of lexicographic practice across a variety of dictionary cultures, so there is no reason why it should not feature prominently in electronic dictionaries. But there are new considerations that follow from the nature of electronic dictionaries.

In English pedagogical lexicography definitions are now usually written using a controlled defining vocabulary of about three thousand (give or take a few hundred) items (for a detailed historical account see Cowie 1999). The chief advantage of vocabulary control is that definitions are easier to understand by foreign learners, although there may be unexpected benefits to native speakers as well (McCreary and Amacker 2006). But there are disadvantages, too; one of them is the following: if you restrict your choice of words you can use in the definition to the most common words in their most common meanings, then you lose some of the power to distinguish between finer shades of meaning. In fact, of those learners’ dictionaries that are in principle dedicated to the use of restricted defining vocabulary, most (if not all) find it necessary sometimes to use in the definition a word which is not part of the defining vocabulary. If this happens, the outsider status of such an item is usually indicated with special typography, commonly lower capitals, such as CENTRAL BANK and INTEREST in the following entry from LDOCE4 online, where contrastive font colour is also used:

Federal Reserve Bank, the also the Fed informal the FRB the US CENTRAL BANK, which is divided into twelve banks, each operating in a different area of the US according to the Federal Reserve System. The 'Fed' has an important influence on US economic policy, because it fixes the rate of INTEREST that banks must pay when they borrow money.

The relevant sense will sometimes (though not in our example) be indicated with a raised number if the capitalized term has a polysemous entry. The rationale behind these special efforts is that a substantial proportion of users will have a need to look up the capitalized terms in the dictionary, these being outside the basic vocabulary stock of English. What needs to be taken into account in the design of online learners’ dictionaries is that looking up takes a whole new meaning in electronic dictionaries: laborious page-turning and letter hunting can be replaced with a single mouse click or even hovering your mouse over the target, whereupon a small popup window can display an instant explanation. In our example, the web-based version of LDOCE4\textsuperscript{3}, this potential is not fully realized as of this writing.

\textsuperscript{3} The above comments refer to the new version of the web-based interface, put up in the recent months according to my observations (some time in autumn 2008)
(December 2008): not only does clicking on \textit{interest} fail to take the user to the relevant financial sense right away, but also \textit{central bank} is not recognized as a lexicographic unit; that is, clicking on \textit{central} takes us to the entry for \textit{central}, while clicking on \textit{bank} throws up the entry for \textit{bank}, leaving it to the user to further hunt down the multi-word unit \textit{central bank}, which can ultimately be accessed from both these entries, but only with a bit of scrolling. Obviously, this should not be so: the implicit cross reference is logically being made to the autonomous unit \textit{central bank}, and this is where the user should be taken right away.

To make a more general point, though, given the relative ease with which instantaneous assistance can be given (in principle at least, if not always in practice) for problematic words in the definition, it might make good sense to use words outside the strict defining set more liberally in an online learner’s dictionary than a paper one; or, possibly, to expand the defining vocabulary. Of course, moderation needs to be exercised, as too many infrequent words may negatively impact the readability and comprehension of the definitions. I will come back to the problems of dictionary definitions and how to help overcome them in section 3.2 below.

3.1.2. Equivalent

Equivalents are to bilingual dictionaries what definitions are to monolinguals: the most common mode of meaning provision in this dictionary type (see Adamska-Sałaciak 2006 for a detailed discussion of equivalents in bilingual dictionaries). Although the electronic medium will probably not revolutionize the lexicographers’ approach to dictionary equivalents, \textit{L1}→\textit{L2} electronic dictionaries can enhance the usability of entries for text production and language study by providing instant access from the target language items to more complete information about these items (such as details on their meaning). Recent research into the effectiveness of such extended treatment offers encouraging results (Laufer and Levitzky-Aviad 2006).

However, finding out meaning is not the focal component of foreign language text production, an activity which essentially proceeds from (native-language-encoded) concepts to foreign language form, and so meaning indication is primarily relevant in bilingual dictionaries going from \textit{L2} to \textit{L1}. Here, equivalents remain the most important carrier of meaning (Laufer and Hill 2000; Lew and Doroszewska in preparation), just as they do for paper dictionaries (Lew 2004).

3.1.3. Example

The dictionary example performs a number of functions (Fox 1987; Purczyńska 2002; Toope 1996), including some not very obvious ones, such as guidance on grammar (Bogaards and Van der Kloot 2001, 2002). However, in this context the contribution of exemplification to the explanation of meaning is most relevant, and in this connection it is worth pointing out the synergistic effect of the definition and example which has emerged from the relevant studies (Summers 1988; Laufer 1993).

In recent decades, a debate has been going on about the role of corpora in providing examples. The three broad approaches are corpus-derived (examples taken out of corpora with minimal modification), corpus-based (adapted from corpus material), and invented, with arguments being presented for — and against — each of the above (Cowie 1989; Humblé 1998, 2001; Fox 1987; Laufer 1992). In the current versions of English monolingual learners’ dictionaries, corpus-based examples seem to enjoy the greatest popularity, being perceived as something of a compromise between raw, unedited citations and invented examples.

Text corpora can be bundled with the electronic dictionary, serving as an on-line source of additional examples. Examples can then be available at a click from relevant points.
in an entry, as a concordance generated on the fly from the corpus, thus partially solving the problem of space putting a restriction on how much exemplification a dictionary can provide (Xu 2008). As early as 1995, COBUILD2 in its CD-ROM version included a (relatively small) 5-million-word sample from the WordBank corpus. However, the integration between the main text of the dictionary and the corpus sample left much to be desired.

In this connection, the GDEX tool (Kilgarriff et al. 2008), an automated procedure capable of ranking corpus citations, holds great promise in giving dictionary users access to useful corpus-derived examples without being flooded with too much irrelevant material.

3.2. Audio presentation

In an electronic dictionary with audio capability, there is a possibility of rendering the verbal components not just in traditional spelling, but as the spoken word, either in the form of voice recordings from a human reader, or as synthesized speech. Not uncommonly, this option is already implemented for the headword itself, but users could potentially benefit from audio representation of other macrostructural elements, notably the definition and examples.

In section 3.1.1 above I have mentioned problems with simplified definitions in learners’ dictionaries. Yet another disadvantage of using restricted defining vocabulary is the sad fact that you pay off in complex grammar and verbosity what you gain in simplifying lexis – and this can lead to parsing problems, especially by readers who are learners of the language, and thus have imperfect command of it. However, if definitions are read out aloud as well as displayed, then learners will be helped in this task by cues provided by prosody, stress, rhythm, and intonation. Not to mention the long-term benefits of being exposed to authentic pronunciation. The same rationale might apply to examples; in fact, LDOCE4 in its CD-ROM version already provides digitized audio recordings for examples (these are available in the free online version for words starting with the letters D and S).

In addition, audio presentation would give the much-talked-about exposure to ‘real language’ – i.e. speech, also serving as a pronunciation model. If we welcome the tendency for EMLD to make definitions more like teachers speaking to students – well, here’s how they can really be made to speak.

To what extent (if at all) language learners benefit from audio versions of definitions and examples will hopefully be revealed in future user studies. However, we do have some positive evidence with regard to the verbal presentation of the headword for Chinese learners of English: Laufer and Hill (2000) compared the preferences for meaning representation options of Israeli and Chinese (Hong Kong) university-level learners of English, using an experimental electronic dictionary interface. Apart from the traditional microstructural elements: L1 translation, L2 definition, illustrative examples, etymology, ‘extra’ information (other forms of the word, phonemic transcription, register, complementation, related meanings, and other semantic and syntactic details; Laufer and Hill 2000: 61), the experimental dictionary included spoken recordings of headwords. Interestingly, the Chinese students made use of the recorded pronunciations much more frequently than the Israeli subjects. One possible reason for this might be that ‘Chinese preference for the pronunciation option could somehow be related to the fact that Chinese dictionaries are arranged according to the phonetic radical and so Chinese lookup (sic) words in a dictionary by sound’ (Laufer and Hill 2000: 70).

3.3. Non-verbal representation

Verbal indication of meaning has been the way of lexicographic practice for centuries, and it remains just as important today; but pictorial illustrations are also well attested in paper lexicography. This, plus some other electronic-only options (not possible on printed paper) will be discussed below.
3.3.1. Pictorial illustration (static)

According to the information in the front matter of LDOCE2 (F49), in this learner’s dictionary four groups of items are supported with illustrations:

1. common animals, plants, objects;
2. things not easily explained in words, such as shapes, complex actions;
3. groups of related words;
4. basic meaning of words often used in a figurative or abstract way.

A more elaborate classification of items explained through illustration (based on an analysis of several dictionaries, including two English learners’ dictionaries) is offered by Stein (1991). So much for the categories of items supported through pictures. But what about their effectiveness?

A study by Nesi (1998) looked at the success in conveying the meaning of everyday household objects with readily recognizable visual features (the target words were: colander, insole, plunger, shoehorn and spout). Nesi found that even in those cases where subjects (mostly undergraduate students) were misled by the definitions and illustrative examples (as evidenced by the sentences they produced when prompted), many of them were subsequently able to correctly identify the objects among a set of twelve simple pictures in the form of black-and-white line drawings. This would suggest that there is value in enhancing verbal explanation of meaning with pictorial illustrations.

More recently, Gumkowska (2008) studied vocabulary retention by Polish learners of English from bilingual-dictionary-like concrete noun entries with and without pictures. Using a cross-balanced design, she found the immediate retention rate for the picture-enhanced headwords to be around 80%, compared with 70% for the entries with bare Polish translations. This difference turned out to be significant despite a small sample (N=20, \( t_{paired} = 2.58, df = 38, p = 1.4\% \)).

The two studies by Nesi and Gumkowska point to the dual benefits of pictorial illustrations as an additional meaning indicator in dictionaries, both for immediate recognition/comprehension of concrete noun vocabulary items, and for vocabulary acquisition.

Clearly, given the virtually unlimited storage space of most electronic dictionaries (handhelds possibly excluded, at least for the moment; compare the discussion in section 2.2 above), being able to include pictorial illustration for a greater number of entries and senses than has been standard in paper dictionaries is an attractive option. However, there is still the problem of restricted presentation space, quite severely restricted in the case of handheld devices, including dictionaries on mobile phones: it is at present technically impossible to reconcile portability with a comfortably large display.

Another difference vis-à-vis paper dictionaries is that the use of colour is very expensive in print, while colour displays are now standard except on some types of portable handheld devices. This means that colour illustrations can be used more readily in principle, although we are lacking evidence at the moment to demonstrate that colour illustrations are any more effective than the more traditional simple iconic line drawings.

---

4 Ilson (1987), on the other hand, uses a broader concept of dictionary illustration which also includes tables and diagrams
5 This situation could change in the near future with the introduction of projector glasses or, in the not-so-near future, 3D hologram projectors
3.3.2. **Photographs**

Photographs may in many cases be easier to obtain than drawings (for readily accessible real everyday objects, that is; a photograph of Bigfoot would be rather hard to obtain!) and can be included in dictionaries, electronic more easily than paper, especially on devices with colour screens. However, one disadvantage of photographs is that, unlike a hand drawing, they do not emphasize what is the focus, i.e. which part of the pictorial illustration represents the object (and lexical item) in question. In a hand drawing, the respective elements can be foregrounded to guide the viewer in this regard. For example, to properly represent a **thumb** in an illustration, it has to be represented as part of the hand; at the same time, it has to be indicated that the lexical item refers to just the thumb, and not the palm or any other fingers. A skilled artist can then use a variety of techniques (broken lines, shading) to foreground and background fragments of the picture as appropriate. Figure 2 below illustrates this for the item **fin**. However, manipulating a photograph to achieve a similar effect, such as adding an arrow pointing at the thumb, is also conceivable and has been used in lexicography.

![Figure 2: An illustrative picture for the entry **fin** showing how shading and the use of an arrow can help identify the defined item in line drawings, so that it is not taken to signify the holonym of which it forms part (after Gumkowska 2008)](image)

3.3.3. **Graphs**

Iconic illustrations and photographs may work well for representing every-day objects, but what about words expressing more abstract relationships, such as prepositions? Adamska-Sałaciak (2008), arguing for a lexicography informed by cognitive linguistics, advocates the inclusion of schematic graphs to represent the meaning of prepositions, such as the one for the English preposition **over** represented in Figure 3.

![Figure 3: A graph proposed for inclusion in bilingual dictionaries as a representation of the proto-scene for the English preposition **over** (Adamska-Sałaciak 2008: 1482)](image)

It would seem that such schematic illustrations might also hold promise in the representation of various spatial and meaning relations. A good example of how the latter can be shown in an effective way is the Visuwords graphical interface (www.visuwords.com), currently serving to visually represent the lexical relationships data from WordNet, but it
could be used as a front-end for other lexical databases. An example entry for the lemma disaster is shown in Figure 3 below.

![disaster graph](http://www.visuwords.com/?word=disaster)

Figure 3: A floating quasi-3D graph illustrating sense relations for the entry disaster in the Visuwords interface (http://www.visuwords.com/?word=disaster)

3.3.4. Animation

Animation seems well suited to a selection of Stein’s (1991: 109-111) categories, such as countable nouns shown in ‘different sides or aspects … to capture different dimensions or stages of use’, an example being a picture of concertina; and even more so category 3, ‘countable nouns that denote an event or an activity that is realized by a number of event or action phases’, examples of which are eclipse and press-up/push up. Then there are action verbs such as dive, inflate (a balloon), mushroom, inject, erupt and bend. In entry types such as the above, it would seem that an animated illustration would provide a more user-friendly representation of stages or progression of an action than a static drawing.

That is also what Lew and Doroszewska (in preparation) expected to find in their study of the effectiveness of animations in on-line dictionary entries. Quite surprisingly, consulting animations not only did not seem to help Polish students, but was found to negatively affect item retention, perhaps because it distracted the subjects from the relationship between linguistic form and meaning. The items studied were verbs describing body actions such as shiver, blush and frown. In view of this outcome, I would not recommend the inclusion of too many animations in future electronic dictionaries before more evidence is available on the effectiveness of animations for different types of entries.
3.3.5. Video sequences

Video sequences are still rather costly in terms of storage space; they are also rather expensive to make, and their consultation can not be very quick (often an important factor in dictionary use). Their greatest usefulness is to be expected for highly conventionalized exchanges. Through the detail available in videos, information on pragmatic/situational context can be conveyed. Some video sequences are already available in learners’ dictionaries (LDOCE4, for example).

4. Closing comments

Given the technical potential of modern electronic dictionaries, various possibilities present themselves of combining the different techniques of meaning presentation in electronic dictionaries. Multi-modal processing should, in theory, engage the learner more and bring about improved comprehension and learning, but in practice this is not always so (Lew and Doroszewska in preparation). Clearly, much more user research is needed before we are able to find the optimal (combination of) ways of presenting meaning to learners. It should not be surprising if the solutions turn out to be sensitive to factors such as: dictionary culture, consultation goals and context, level of L2 proficiency, type of lexical item.

References

Dictionaries


Other references


Lew, R. and Doroszewska, J. in preparation. ‘The Effectiveness of Animation as a Representation of Body Action Verbs in Electronic Dictionaries.’


