Dictionary Entry and Access
Trying to see Trees and Woods

Geart van der Meer
University of Groningen
Department of English Language and Culture
PO Box 716
9700AS Groningen
The Netherlands
vdmeer@let.rug.nl

Abstract

In this paper an attempt is made to combine the needs of quick information retrieval in a monolingual (English, and in particular an English learner’s) dictionary with the ideal of vocabulary development by means of the dictionary. To this end the author gives a sample entry of the word cut (noun and verb) and draws conclusions about the feasibility of this combination.

1 Introduction

In the modern learner’s dictionaries one can observe a battle between the requirements of quick access and decoding and a desire to place language facts in a wider context, so that one not only increases one’s knowledge about the problem originally triggering the search but also learns additional facts that will come in handy in future active use of the language. Decoding generally means solving one problem at a time (for instance what does a word mean in the context where I find it, i.e. receptive meaning retrieval). However, productive vocabulary development has also always been one of the declared or implied aims of learner’s dictionaries. In other words, in addition to being a passive monolingual dictionary, a learner’s dictionary should also be an active tool enabling the study of the language’s vocabulary in a meaningfully guided and structured way. The latter means that the information should be presented in such a way that it coherently and systematically shows links between meanings, between words (both from a morphological and semantic point of view), between multi-word lexical items and the key words therein, and generally prevents as far as is reasonable the atomisation resulting from the frequency-based principle of sense ordering (as applied in the modern learner’s dictionaries).

Furthermore, very clear semantic, though not formal, links between words like moon – lunar, sun – solar, dog – canine, cat – feline are not pointed out systematically in any of the ‘big four’ [Van der Meer 2001], though it would obviously be very helpful for the foreign learner to be told at sun or sunny that it would be profitable to have a look at solar as well. Clearly, all this is far from ideal if one holds that the learner’s dictionary should also encourage the user to actively explore links and patterns in the English vocabulary in order to learn to understand and use it in a near-native way.
In this paper I will explore whether it is possible to combine the requirements of *quick retrieval* (the quick problem solution approach) with the requirements of *vocabulary building*. By quick retrieval I will understand being able to find the required information without necessarily being force-fed other, possibly at that moment irrelevant, information first, whereas vocabulary building means being shown links between meanings, patterns of meaning and links between morphologically and semantically related words while being guided to the information desired. Up to now, the former has generally led to a fragmented presentation of seemingly independent senses. The latter (as attempted on a large scale in CIDE\textsuperscript{5} or more modestly in NODE\textsuperscript{6}) may cause retrieval problems, while at the same time it has been noted that many decisions (as by the CIDE editors) regarding which lexical items (i.e. usually derivatives and compounds) should semantically be linked with which other items, were highly debatable. This may already serve as a caveat: links between meanings and lexical items will often be vague and hard to describe. One of the main causes of this difficulty is the fact that many words have long histories in which frequently connections may be lost or obscured.

2 The problem: atomisation or integration?

As already stated, the atomisation of presentation as observable in *inter alia* the English learner's dictionaries is the consequence of giving priority to frequency of occurrence of senses: the most frequent sense is discussed first, regardless of whether this may be a case of, for instance, metonymy or metaphor that is clearly derived from a still current literal, or 'basic', sense. This method has received a tremendous boost from the availability of huge digitised corpora, for the first time giving lexicographers an indication of which sense is the most often used. Though chances are that the most frequent sense will already be known to learners, it is equally true that this is the sense they will most often come across and will hence be able to find most quickly when necessary. However, I doubt how well served learners really are by a method which often fails to encourage them to explore the more basic, though perhaps less frequent, senses of words – the 'basic sense' being that still current sense from which other senses can readily be derived by means of such well-known processes as metonymy, metaphor, generalisation, specialisation and synecdoche, amelioration and pejoration, and *not in any acceptable way the other way round*. It is my firm belief that a native speaker 'knows' that the two current senses of for example *underbelly* are related as respectively ‘basic’ and ‘derived’ (as an established metaphor). Generally, the frequency-based method of presentation impedes exploring and learning meanings coherently and in all their ramifications. In short, it blocks an *integrated* view of meaning.

It is unrealistic to believe that the meaning of words is stored in the brain on the pattern of dictionaries, with their two-dimensional linear presentation. Yet, it seems likely that some presentations are more ‘realistic’ than others. Thus, the NODE, despite its not always flawless application of its own principles [Van der Meer 2000], is a laudable and courageous attempt to group sense definitions on the so-called *core sense – subsense principle*, whereby the derived senses (subsenses) are placed under the heading of the core, which stands for a more ‘literal’ meaning. Many older dictionaries, compiled before the advent of the huge computerised databases, tended to do something similar, using words like ‘*hence*’ or labels like ‘*fig*’ to indicate that a certain sense was somehow based on or derived from another.
The dictionary-making process

There will be very few or no modern lexicographers who do not admit that the semantics of words has structure and is not a list of unconnected, and numbered, senses. Yet, it seems that, with the help of the digitised databases, those favouring quick retrieval have worsted those favouring overtly presenting semantic links. Let us examine, by studying one particular word (i.e. cut), whether these two approaches cannot be combined in one dictionary entry. My reasoning will be as follows.

Since the requirements of the two approaches are rather different the presentation will have to be designed in such a way that the description of the holistic aspects and the demonstration of the coherence of meaning does not get (too much) in the way of quick access to the particular sense or use. In the first place I propose to let every entry be opened by a semantic synopsis (called here the ‘semantic profile’) describing in a summarising but clear style the major meanings of the entry-word, their coherence and links, such as metaphorically derived established usages, and normally provided with one clear example for each. Thus, this profile enables one, in concise form, to quickly absorb a full picture of the semantics of the entry-word. This should in many cases already be enough information.

In this profile there will be numbering referring to more examples to enable the user (whether or not interested in reading on in the profile itself) to refine his understanding of the particular sense he is looking for. These further examples may be provided with explanatory glosses and will, when applicable, move from more concrete to more derived or figurative uses, while important word combinations may be highlighted. In the particular case of cut it proved necessary to continue with a section with an alphabetical list of fixed combinations (e.g. cut out) that need further elaboration by means of exemplification and explanatory glosses. Ideally, no genuine sense definitions need be given here since they all still fit under the semantic umbrella of the profile. In a final section uses which cannot, or which can only with some difficulty, be linked with the profile are dealt with. If semantic links are not clear at all, there is an obvious case for adopting an approach in terms of homonyms.

We thus have the following main outline of an entry as for cut:

<table>
<thead>
<tr>
<th>1. SEMANTIC PROFILE</th>
<th>with holistic survey of meaning</th>
<th>with a restricted number of examples with relevant context relevant contexts given</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. FURTHER EXAMPLES</td>
<td>illustrating further applications of profile meanings</td>
<td></td>
</tr>
<tr>
<td>(directly linked to profile)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. SECTION WITH SETTLED COMBINATIONS</td>
<td>still linked to profile meanings</td>
<td>with context words in alphabetical order</td>
</tr>
<tr>
<td>4. SECTION WITH RESIDUAL CASES</td>
<td>semantic link with profile is unclear or tenuous</td>
<td>context words (if any) in alphabetical order</td>
</tr>
</tbody>
</table>

The entry itself then looks like:

| 511 |
CUT\textsuperscript{9,10}

- **SEMANTIC PROFILE** \textsuperscript{[1][2]}

**verb:**

1. you use *cut* as a verb\textsuperscript{11} when a sharp object (typically a knife or scissors, or a similar object) divides, or is used to divide something, either partly or completely (cut a ribbon / a string); cut may also be used figuratively \textsuperscript{[4]} in most senses described here (his cruel remarks cut her deeply (= hurt emotionally));\textsuperscript{[5]} you also use *cut* when you remove material: cut a page from/out of a book; (fig.) this scene was cut from/out of the final version of the film;\textsuperscript{[7]}

2. the result of the action of dividing may also be mentioned (the thieves cut a hole in the fence; he was cutting himself a piece of cake);

3. the action of dividing may lead to a reduction in size of one part of an object or an activity (cut the grass; (fig.) Branagh has cut the play judiciously);

4. you may also use *cut* in a figurative sense when an activity is ended or stopped: (cut food and water supplies);

**noun:** \textsuperscript{[4]}

5. you use *cut* as a noun to refer to the action of cutting or its result (also frequently figurative): cuts and bruises on the face; make a small cut in the material;

**MORE EXAMPLES** (including numerous figurative ones):

1. cut the tomatoes in half vertically; this knife does not cut very well; you can hear the saw as it cuts through the bones; thinly cut cucumber sandwiches; he cut himself shaving; I cut my finger; blood from his cut lip trickled from his chin; Zoe was badly cut as she scrambled down the rocks to reach him; she fell and cut her head open; her face was cut to pieces/shreds by flying glass; two survivors were cut free/loose after being trapped for twenty minutes; I’ve cut my hand on that glass; long canoes cutting through the waves; (of clothes) badly cut (= designed and made) blue suits; cut a pack of playing cards (= divide it into two); a deal to cut 50 billion dollars from the federal deficit;

2. *cut a record/album/disc/CD* (= making a recording of music); many infants do not cut their first tooth (= when it grows through the gum) until they are a year old;\textsuperscript{12} also see teeth;

3. cut your toenails; you have had your hair cut; the first priority is to cut costs/prices/losses/wages etc;

4. cut classes/lessons/school (= not go to classes or school); cut the euphemisms, Brenda snapped; cut an engine / a motor (= switch off);

5. the operation involves making several cuts in the cornea; a cut on his left eyebrow; price/tax/job cuts; a cut in interest rates; a cut of meat; a lean cut of pork; cuts in public spending / living standards; we’ve had to make some cuts in the text; cuts in electricity and water supplies; lawyers take their cut (= share) of the little guy’s winnings; starting at £17 for a cut (= a hair-cut) and blow-dry; the elegant cut of her dress;

*cut* (verb or noun) **IN CONTEXT** (including numerous figurative uses): \textsuperscript{[10] [11]}

- above: Her detective stories are a cut above (= better than) the rest;

- across/through: he decided to cut across/through the heath (= take the shortest distance); the problem cuts across (= affects or is important to more than one) all party lines;

- away/out ↔: cut away/out dead wood;

- back: cut back ↔ the rose bushes (= prune); cut back (= reduce) expenditure, cut back on defence spending; cut back on alcohol; they suddenly cut back and headed north (= suddenly went in the other direction);

- both/two: this publicity cuts both/two ways (= has two possibly opposite effects);
• **down** ↔: cut down a tree; cut down costs (= reduce); cut down atmospheric pollution; we must cut down (= shorten) the article to 2,000 words; cut down to size, cf. size; cut down on coffee and cigarettes;

• **from** (film, television etc.) the scene cuts from (= moves quickly from) the bedroom to the street;

• **in** ('Forget it,' she cut in (= interrupted); she kept cutting in on our conversation; (of traffic or people) impatient drivers often cut in (= move or push suddenly and dangerously in front of you); (of engine) emergency generators cut in (= started working); cf. cut out;

• **into** Saturday shopping cuts into my weekend (= makes it shorter);

• **off** ↔: he cut off a generous piece of meat; they cut off (= blocked) the enemy's retreat; we got cut off by the incoming tide; our water supply / our phone has been cut off; we were cut off (= the telephone line was disconnected) in the middle of our conversation; he cut his son off (= disinherited him) without a penny; the exiles had been cut off from all contact with their homeland;

• **out** ↔: cut out the coupon and send those cheques off today; cut sth out of a newspaper; I would cut out the bit about working as a waitress; he has cut out a niche for himself; the new administration has its work cut out for it (= it is a difficult job); cut that behaviour out (= stop it); cut it out! (= stop it); a guilty plea cuts out (= removes) the need for along trial; since my heart attack I've cut fatty foods out (= stopped eating) altogether; cut me out! (= I won't join you); these overhanging branches cut out (= block) the sunlight; do not cut out when everybody is travelling fast (= suddenly leave a line of traffic); this would cut them out of the debate over what to do with public lands; he is not cut out for teaching / cut out to be (= does not have the qualities and abilities) a teacher; (of engines etc.) the helicopter crash landed when one of its two engines cut out (= stopped working); the heating cuts out automatically;

• **through**: cut through, see across; also: They use a machete to cut through the bush; she always manages to cut through (= understand) a complex theory and get at the facts;

• **up** ↔: cut up the meat; he was badly cut up (= injured) in the fight; she was pretty cut up (= upset) about them leaving; he moved from lane to lane, cutting everyone up (= moving too close in front of);

- **other or special uses:** "12" cut!, stop filming (command by director)

- **cut and dried** (= clear and definite);

- **cut and paste**: move sth. to another place on a computer screen;

- his cowardly decision to cut and run (= escape from a difficult situation);

- the cut and thrust of parliamentary debate (= the lively or aggressive way sth is done);

- also see cackle, corner(s), crap, dead, figure, film, fine, ground, heart, ice, knife, loose, mustard, nose, penny, quick, rough, story, short, swath, will, penny, teeth, will [15][14]
3 Remarks

In this section a number of remarks will be made further explaining the rationale behind the proposal in section 2. The numbers here correspond to the superscript numerals in square brackets \(^\{1}\) in the cut entry above.

1. The semantic profile: the semantic profile should be written in such a way that the average user can grasp the main outline of the meaning of an entry-word and will in many cases not really need to continue his search if his primary aim is merely to understand the meaning of a word. Ideally, any particular use and particular combination should be interpretable (at least in its main outlines) with the aid of the profile. In case this does not succeed or more details are required, access from the profile to the section offering more examples should be as straightforward as possible by means of numbering. Should this still not be enough the third section (with fixed combinations) will have to be accessed, but then the profile’s role has come to an end. Experienced users may often guess that this is necessary and may then skip the first two sections.

   Should users wish to use the entry-word productively, they may of course likewise have to continue their search, and the profile design should be such that access to further information is made possible directly from the profile. Failing this, the section with fixed combinations should in this case be consulted as well.

   The general philosophy behind the profile is the reasoning that multiple meanings of semantically complex words often form a coherent whole or at the very least clusters of coherent wholes, where most relations between various meanings (or uses)\(^{13}\) can be described in the well-known terms of metonymic or metaphorical extension, and also generalisation, specialisation and synecdoche, amelioration and pejoration. My expectation is that many lexical items will allow this approach. However, it should also be anticipated that there will be numerous cases where there is a vague link between the multiple meanings, which is nevertheless extremely hard to fit into the above categories.\(^{14}\) Language, it should be realised, is the product of centuries of development and this fact will often produce obscure links between meanings, either because certain missing links have been lost in the course of time or because we are simply no longer able to grasp the ‘logic’ behind some sense developments. Thus, a certain number of residual cases resisting attempts at analysis must be taken into account.

   Yet, all this does not remove the fact that it is often quite well possible to present the semantics of complex items in a much more coherent, integrated and hence much more intelligible way than in current dictionaries with frequency-based sense ordering. In the final analysis, frequency is of course a fact of language and presumably acknowledged to some extent by linguistic intuition, but it can say nothing about semantic links and patterns. The only thing a frequency-based order does is increase look-up speed for isolated senses. No more.

2. In a normal learner’s dictionary there should be grammar codes. I have left them out here for ease of exposition, since their absence will not affect the general argument.

3. To compensate for the loss of ‘catchwords’ (cf. next section) to guide the user to a quick ‘hit’ it is perhaps possible to use small capitals or a similar device to highlight the most important elements in a definition.

4. Or should I write ‘non-literally’? Clearly, the semantic profile approach to the problem of sense definition confronts us again with the age-old problem of figurative or non-literal
language, with what ‘literal’ (‘figurative’ or ‘basic’) means and generally with how to
deal with extensions of meaning and in which order to present them. It seems obvious to
me that in dictionaries describing contemporary language the profile should reflect the
synchrony of language and not its diachrony. Since I have also rejected the frequency-
based approach to the organisation of the entry, I am (in Cruse’s terms) left with three
possibilities as to which sense to select as ‘basic’ or ‘literal’ [Cruse 2000: 199-200]: the
default reading, the reading from which the most plausible path of change begins, and the
reading most closely related to basic human experience. The first refers to the sense that
comes to mind when the word is considered out of context; the second to the sense from
which one can most plausibly derive other senses by means of the familiar meaning-
extending process referred to above. The case of cut does not seem to pose too many
problems, but Cruse points to the fact that in other cases opinions may differ. The third
reading reflects the fact that much of our more abstract vocabulary is ultimately based on
words describing our unmediated cognition through our senses. Thus, see as ‘visual
experience’ is basic as direct (visual) cognition, whereas the ‘understand’ sense is not.\[15\]
If a dictionary, on the principles as sketched here in general terms, is ever going to be
compiled a combination of these three seems to me to be called for. The strongest case is
when all three apply. It is worth repeating that diachronic reasoning should not bias our
choices: a supposedly ‘basic’ or ‘literal’ sense should be a clear synchronic fact. If not, it
has to be rejected. Thus, as long as the literal readings of for example defuse and morass
are still in use and may be assumed to be known to the average user they should be taken
as basic, though they may be less frequent than their metaphorical readings [cf. Van der
Meer 1996; 1997].

Perhaps a note of warning is in order here: though my proposal is to write the profile in
the manner indicated here, it should not degenerate into a theoretical exercise. The aim
should firmly remain explaining to the user in as simple a manner as possible the full
meaning of a lexical item and if in that full meaning there are tenuous links or even
unexplainable elements, these should be presented as such: the lexicographer should not
pretend to offer more than he can realistically demonstrate.

5. The so-called explanatory glosses should be chosen with the utmost care, in order to
prevent them from being taken for sense definitions proper. A gloss represents this
particular sense in this particular context and need not be relevant in a different context.
Especially in the case of figurative uses the glosses may differ per context even though
essentially the same metaphor is used (cf. for instance the glosses for defuse and morass
under 4. in the examples section.

6. To the extent that this is feasible, attempts should be made to use examples in such a
way that typical and frequent combinations (collocations etc.) are made prominent. I have
done so here by bold-facing (also cf. deeply a couple of lines earlier). This is in line with
the practice of for example CIDE.

7. It should be understood that the numbering of the sense definitions has been adopted on
purely practical grounds. In order to keep the access structure manageable (especially the
move from the profile to the examples section) a certain amount of numbering is
unavoidable. Yet, the entire reasoning behind the profile, with its stress on the coherence
of the senses and their interlinking and interdependent nature, should make it absolutely
clear that such numbering has no deeper meaning and does not suggest that senses can be neatly cut up and counted.

8. The semantic link between the noun and the verb is relatively straightforward here and can hence easily be accommodated in one profile, but this may not always be so. Though the structure of the entries should be standardised as much as is practicable, it should be borne in mind that it may not always be possible to combine both the verb and the noun readings\(^{16}\) in one profile. The aim should always be to combine in one profile whatever can reasonably be shown to belong to one integrated and coherent ‘semantic picture’. Any reading not accommodatable in this way should be dealt with singly or as belonging to a group of other cases, as the case may be. The problem I am here referring to is, of course, the choice between distinguishing between separate profiles of one lexical item or between two homonyms. My feeling is that the profile format should make us favour the homonym approach. This would have to be tested more fully, but it would seem that two profiles for one lexical item would be an inconsistency given the above explanation of the term ‘profile’, which ensures that all readings in a profile must form a coherent whole.\(^{17}\)

9. I suggest that here the order of the examples should be rather from more literal to more figurative, mirroring the order in the profile; hence it is not advisable to apply any rigid alphabetical order.

10. I have decided to restrict this list in such a way that in the relevant contexts only the ‘minor’ grammatical classes (such as prepositions, adverbs, pronouns) are admitted. The rest will have to be cross-referred to (cf. the section under ‘\(\text{also see}\)’). This is, however, a decision each editorial team will have to make for itself.

11. The semantics of these combinations should, with the support of the examples themselves, still be derivable from the general definitions in the profile. If not, they should go to the next section (‘\(\text{other or special uses}\)’) or even be placed under a homonym.

12. This is, admittedly, a little like a ragbag of leftovers. All of them, with the exception of the first example, might just as well be accommodated under the next section, where the user is sent to entries elsewhere. I have left these cases here to illustrate the problem facing the editor. It will depend on the policy of the particular dictionary which word-class takes ‘consultation precedence’ over which other word-class in the case of phrases, collocations, idioms and the like. One might think of such strategies as Van Dale’s (former?) 1-2-3 principle, which means that the search path is ‘go the first noun, then - failing this – the first adjective and then - failing this also-, the first verb when reading from left to right. It will be clear that this is often less simple than it looks, given the possible variability of word-order,\(^{18}\) and requires a certain expertise in grammatical analysis, which users may not always be prepared or able to make.

13. See the point raised under 11.

14. A problem I have evaded so far is that of the morphologically related words (e.g. derivatives like \textit{cutting}) and compounds (like \textit{cutting edge}) forming so-called word families. The principle of vocabulary development mentioned earlier might be interpreted as favouring so-called nesting of such items. This method would gather all morphemes with the same or similar meaning in one complex entry, treating derivatives and possibly also compounds as ‘run-ons’ without full headword status, which would hence facilitate this vocabulary development, and would in many cases also make sense definitions easier, since repetitions of identical semantic information may be avoided. Yet, many
dictionaries seem to be moving away from nesting (cf. the COD as from 1995 and the OALD6). Though nesting has some obvious advantages they are outweighed by disadvantages in that the search path becomes more difficult, since the alphabetical order in the macrostructure may be disrupted and the entries may become too cluttered and complicated. For learner’s dictionaries perhaps occasional cross-references will be useful to make users aware of links with derivatives and compounds.

4 Conclusions
The above sample entry seems to demonstrate that the aims of quick information retrieval and vocabulary development are not completely irreconcilable within one and the same entry. Yet, this is only one sample and more would have to be written to more reliably test the viability of this project. In this particular case my impression is that quick information retrieval is less well served than vocabulary development in its broadest interpretation: the recently developed search path shortening device of the ‘catchword’ (i.e. variously called guideword (CIDE), short cut (OALD6) or signpost (LDOCE3)) is certainly a great help in scoring a ‘hit’ straightaway and will have to be abandoned in my proposal. There will in my view, however, be a huge gain in what might be termed ‘integrated semantic awareness’, a firm grasp of the entire meaning of an entry and its possible applications. I also feel that, if we can manage to hone the semantic profile technique to greater perfection, this will also result in a lot of space being saved. A well-written semantic synopsis will make a lot of too fine distinctions superfluous since they naturally follow from the profile definition. For the learner who has advanced beyond the beginner’s stage the profile plus some well-chosen examples will suffice to draw reliable conclusions about semantic possibilities, while for more active use the section with more examples will have to present further and more detailed suggestions, both in terms of the further use of the entry word and its immediate surroundings (syntactic information and collocational choices).

Endnotes
1 For instance by pointing out that metonymy and metaphor are based on a core meaning.
2 For instance by pointing out morphological relationships and semantic similarity (synonyms).
3 For instance by not obscuring, as in LDOCE3, the fact that pull out all the stops is linked to stop as a musical term.
4 It seems obvious to me that the curiously mixed descent of English, with its large Latin/French-based vocabulary and the learning difficulties this causes, should receive the attention it deserves.
5 That is, especially its heavy nesting of derivatives and compounds.
6 Cf. Van der Meer [2000] for NODE.
7 Or simply ‘profile’. The idea for this term was suggested by the (Dutch) Van Dale bilingual dictionaries, which use translation profile for the section of the entry before the diamond (♦) where context-free translations are presented.
8 Cf. the discussion below about the choice between a residue as envisaged here and treating hard cases under another entry, i.e. a homonym.
9 The superscript numbers in [..] refer to the remarks following this section.
10 The following is a concoction based on OALD6, CC3, CIDF and LDOCE3.
11 I here opted for the full-sentence definition type. Strictly speaking, this choice is unrelated to the choice of the organisational principle of the entry. It is up to the policy of the dictionary which defining style to choose, independently of which microstructure its editors prefer.
I admit that the metaphor is here not very clear and that hence there might be a case for moving this expression elsewhere.

It may in some cases be argued that an established metaphorical use of a lexical item is in fact no more than that and not a separate sense.

Take for example the word *clear* (adj.) as I tried to describe it lexicographically in Van der Meer [2000].


Or whatever possible combinations of grammatical functions occur.

As argued in Van der Meer [2000] the NODE practice of combining various core senses in one entry may be objected to on similar grounds.

Especially in languages like Dutch.

This umbrella term was coined by Heuberger [2000].

References

Dictionaries mentioned:

The Chambers Dictionary, Edinburgh 1998


Other works:


Van der Meer, Geart. 2001. “Nonmorphological Derivations” and the Four Main English Learner's Dictionaries. *Lexikos* 11: 1-11