

## On the Status of Word Combinations in the English Monolingual Learner's Dictionaries: Entry or Collocation?

Geert van der Meer  
Department of English,  
University of Groningen,  
PO Box 716,  
9700AS Groningen  
The Netherlands  
geert.van.der.meer@rug.nl

### Abstract

This paper discusses the boundaries between collocations and dictionary entries proper. It compares the way the *Oxford Collocations Dictionary for Students of English (OCDSE)* (2002) deals with collocations as compared with the monolingual English learner's dictionaries. The conclusion is that for the latter many OCDSE collocations are real entries with headword status. The paper concludes with some tentative conclusions about the most practical procedure to be adopted.

### 1 Introduction

When learning a foreign language one is confronted with a number of problems. Among these, one of the most important but initially often underrated ones is mastering the list of (ready-made or more or less ready-made) independently usable form-meaning pairings (cf. Warren 2004) with which one can make larger meaningful units, or to put it simply: learning the list of directly available meaningful language elements that are there for you to pick up and use without having to construct them yourself. This 'list' obviously includes words, but also combinations of words, like expressions and idioms. At a certain advanced level one often knows the most important words and the most important expressions and idioms, but once arrived at that stage there is also a growing awareness that one frequently does not really know how to *combine* them in a completely natural way, the way the natives speak: I am here of course referring to the phenomenon called *collocations*. Such an awareness is testimony to the fact that one is beginning to develop a certain feeling for the language from within, for without this feeling this awareness would not be there. Learning the proper way to combine words is the final phase of mastering a language and while it is perhaps not the most difficult it is certainly the most time-consuming one. Since the number of proper conventional word combinations runs into tens of thousands at least, only a very frequent exposure to the foreign language and of course frequent active use will lead to a native or near-native mastery of the language in this respect. It is my contention that the most conventional collo-

cations belong to the 'list' referred to above and that hence such combinations should ideally all be treated in a dictionary. This paper will examine some aspects of the way in which English MLDs<sup>1</sup> cover the 'list' every fluent speaker has to know, with initially special emphasis on what we call collocations. I will do this by comparing these dictionaries with the OCDSE, a recently published dictionary of collocations.

One of the tools that might assist the foreign learner in the process of learning the list is the dictionary of collocations. In English, the two main ones are the BBI and the more recent OCDSE. A lot has been written about the proper theoretical definition and types of collocation, including by myself (see Van der Meer 1998 and the references there), but in this paper I am not planning to add to this discussion. Instead, I will first informally discuss the concept of collocation as defined and used in OCDSE and then draw practical and perhaps to some extent mildly theoretical conclusions from confronting the OCDSE with the data offered in the five main English MLDs: OALD7, LDOCE4, MED, CALD, COBUILD4.

## 2 The OCDSE and its method

OCDSE itself, obviously, has to answer the question what it purports to describe, namely collocations. In the introduction (vii) it is said that '[c]ollocation is the way words combine in a language to produce natural-sounding speech and writing'. This is, of course, much too vague, so that an example has to clear things up a little: *strong wind* is contrasted with *heavy rain*. The tacit assumption seems here to be that both *strong* and *heavy* are, as it were, concretisations of an *identical* abstract semantic element 'much' (cf. the element *Magn* in Mel'čukian Meaning-Text Theory, as in Steele 1990), but that despite this it is not possible to say *\*strong rain* or *\*heavy wind*. To native speakers the combinations are 'highly predictable', to the learner 'anything but'. Incidentally, the (likewise tacit) assumption that the confusion for foreigners arises from the fact that, in these combinations, both words ultimately 'mean' the same (*Magn*) has to be questioned, on the grounds that *heavy* is much easier to associate with *rain* (water is heavy) and *strong* goes more easily with *wind* (wind can exert a lot of force). In other words, I caution against the sometimes posited arbitrariness of collocations (a point made earlier in Van der Meer 1998 apropos of *perpetrate* and *commit*). This is not to deny the difficulty foreigners experience in this respect, but the causes of this difficulty need not *always* be language-internal arbitrariness of word combinations.

To the extent that genuine arbitrariness *can* be proved, we come dangerously close to the field of idioms, and idioms are excluded from the OCDSE (vii), i.e. completely idiomatic expressions are, but not cases in which for example the collocator – but not the base – is used in a sense rather peculiar to the combination, as in *drive a hard bargain* (viii), where *bargain* is used in its ordinary non-idiomatic sense, but *drive* is arguably used in a sense restricted to this collocation. OCDSE rightly decided that this collocation belonged to its macrostructure, since it is one of the ways in which you speak about *bargain*. Thus, the OCDSE moves be-

---

<sup>1</sup> Monolingual MLDs.

tween the totally free and the totally fixed (*i.a.* idiomatic) combinations (vii) and promises to describe the typical combinations between these extremes (viii). The OCDSE itself does not enlarge upon the concept 'free', but the BBI does (p. xxx): a lexical combination is 'free' if it does not occur 'repeatedly', if its elements are not 'bound specifically to each other' and 'occur with other lexical items freely'. I detect a certain built-in circularity in the first and third conditions, which makes the second the crucial one: yet this condition too is far from compelling, since it is not made very clear what it means if elements are non-free, *i.e.* 'bound specifically' to each other. The only hint the BBI gives is apropos of *commit murder*, where *commit* is 'limited in use to a small number of nouns meaning 'crime', 'wronoing''.

This is the same argument as I above used for the naturalness of the collocation *strong winds*: semantic dovetailing, which makes the combination natural and almost predictable. Yet it does not apply to *e.g. drive a hard bargain* and similar more idiomaticised combinations, for which synchronic semantic naturalness arguments are hard to find. This means that 'bound specifically to each other' may have a number of interpretations.

So it is obvious that the boundaries between 'free', 'collocating' and 'fixed' are fuzzy, but also that within the area of collocations there are several kinds of 'binding'. It is not my intention here to design a definition of collocations that makes it easier to draw the line between the three categories nor do I intend to go deeply into the various degrees and kinds of binding between collocations. Nor does the OCDSE attempt a definition, but it has already been made clear that the collocation as a whole must be non-idiomatic, typical and not completely free. This more in particular implies that collocations are *form-meaning pairings* that are conventional because they 'are associated with a certain salient type of situation or phenomenon' (Warren 2004). In Van der Meer 1998 it was claimed, perhaps less felicitously worded, that such combinations express conventional well-established *complex* (or *composite*) concepts<sup>2</sup> and that therefore the members of the collocation are natural partners semantically, *i.e.* their sense definitions do not clash with, perhaps even presuppose and/or include, each other and certainly fit each other as moving within the same field of discourse or semantic field. Thus, using the example *strong winds*, the full sense definition (and hence the *concept*) of *strong* must include being allowed to apply to phenomena like wind, while conversely *wind* must allow being described as having various degrees of force, and hence inter alia as being *strong*.

Incidentally, the fact that the expected (at least by a foreigner) combination *\*weak wind* as opposed to *strong wind* does not occur according to OCDSE shows that the claim of arbitrariness is after all sometimes true. Instead, we should apparently say for example *light*

---

<sup>2</sup> Klotz (2003: 58) points out that Sinclair's purely statistically based examples like *yesterday's announcement*, *only answer* and *some anxiety* are not in the OCDSE though they are statistically apparently salient. In my view, they are rightly not considered to be genuine collocations, because they do not in any way express established concepts. Admittedly, this is not OCDSE's overt argument and the idea of 'established concept' is a little vague, but the reader will probably agree that *some anxiety* in no way can be said to belong to the same category as for example *strong/high winds*: the concepts expressed in the adjectives *strong/high* 'belong' much more closely to the concept expressed in *winds* than *some* 'belongs' to *anxiety*. Also see the following note.

*wind*, which is interesting: *light* is normally considered the antonym of *heavy*, yet *heavy* is excluded as a collocator for *wind* by OCDSE.<sup>3</sup> This arbitrariness, hence unpredictability, is also found in *high wind(s)*, where *high* is like *drive* in *drive a hard bargain*, used in a sense only found in this combination<sup>4</sup> though *wind* itself has its normal meaning. Thus, between *strong wind* and *high wind(s)* there is a cline of predictability, or one might say, semantic matching: *strong* and *wind* are natural partners given their sense definitions, whereas *high* and *wind* are not and are therefore treated in some dictionaries (LDOCE4, MED) as if their combination<sup>4</sup> were a separate headword like *high street* or *high tea*, independent entries in which *high* has only the most tenuous link with its core senses.

Thus, the justification for including *strong wind* is different from that for *high winds*, since they are 'bound' in rather different ways, and may have something to do with the desire to prevent foreign learners from making native language-based errors, where often very different linguistic combinations may prevail.<sup>5</sup> I suspect that this latter argument may apply to cases which would otherwise qualify for being called 'free' or 'rather free'. As the introduction itself stresses, this book focuses on what are 'typical' combinations.

Some further remarks are in order here about my assumption that genuine collocations are 'ready-made' used to express conventional form-meaning pairings. This idea will have to be clarified in order to prevent confusion with acknowledged multiword lexical items/units, like *joint-stock company*, which may be considered to be a kind of compound and is hence in all five MLDs awarded full lexical headword status. However, OCDSE offers it as a collocation. OCDSE does the same with *joint venture*, which in LDOCE4, under *joint*, has not even been given full headword status, is highlighted in OALD6 in an example sentence and hence presumably considered a collocation, but is awarded full headword status in OALD7, is a full headword in MED, is given in an example under *joint* in CALD but without highlighting – so not even a collocation? – but again found under *venture* in an example sentence as '**joint ventures**' (here with boldfacing of *joint* only!) and in CC3 is again a separate full headword. This example is a wonderful illustration of the apparent uncertainty in English lexicography with respect to the difference between acknowledged *multi-word* lexical items and conventional combinations that we call collocations.

The real problem is the concept of ready-mades or whatever other names might be devised for the various categories of whatever in language is at hand and ready to use, in short everything that is not a 'free' combination but which you use as something *remembered*. The very existence of collocation dictionaries like OCDSE proves that there are considered to be certain ways to combine words in English that are worth learning because they are recurrent:

---

<sup>3</sup> The BBI does give *heavy* as a collocator for *wind*, however! The 50-million free BNC corpus has only five combinations of *heavy wind(s)*, so it is at least not highly frequent.

<sup>4</sup> In LDOCE4 (under *high*) the combination as a whole occurs under 'sense' number 20. A similar treatment we find in MED. Here, no attempt is made to define *high* itself, in isolation and independent of its head noun.

<sup>5</sup> Klotz (2003: 58) also remarks on the fact that semantically quite predictable combinations are included on the grounds of statistical salience as well as 'usefulness to the learner'.

in other words, they are included in the *list of memorised items*: items a speaker uses *without having to actively construct them from scratch, though they may be perfectly regular and analysable*. This means that when one leaves the area of what are called 'free combinations' a huge field opens of ready-made or more or less ready-made building-blocks, word combinations, more or less fixed ways of saying something and so on and so forth.

Within this field there are several categories, the most obvious one of course being the simple word, like *book* or *shop*. Next follow the compounds, like *gift shop*, *bookshop*. The latter, because of its solid spelling, is obviously a 'word', so a lexical item. But *gift shop*, an independent lexical item in e.g. LDOCE4, is listed in OCDSE as a collocation. This is where the confusion already sets in, for *shoe shop*, also an OCDSE collocation, is not even in the LDOCE4 wordlist (it is found together with *shoe polish* under sense 1, but it is not boldfaced and its status is hence rather undefined there) though its status as a settled lexical item is certainly not so very much weaker than *gift shop*.<sup>6</sup> To continue, *shop front* is an LDOCE4 headword, but curiously enough *shop window* is not, while both are OCDSE collocations.<sup>7</sup> Such examples can no doubt be multiplied. What it shows is that the status of English 'words' (i.e. lexical items) versus collocations is far from clear, due *inter alia* to the relative unpredictability of English spelling for compounds (solid, hyphenated or 'spaced' or 'open'). This uncertainty is reflected in OCDSE: solid compounds are not entered as collocations,<sup>8</sup> but a number of lexicalised spaced ones clearly are.

It is, finally, worth remarking that the concept of collocation in OCDSE is used narrowly, in the sense that we are in most cases concerned with immediately *adjacent* words. However, 'collocation' may also more loosely refer to the phenomenon of 'mutual expectancy' or 'habitual association' (cf. Jackson & Amvela 2000: 114, McArthur 1992) of words somewhere in each other's neighbourhood, for example for words like *kettle/water* and *boil*, which do not even have to occur in the same sentence. Here we touch on words within the same 'semantic field' or 'field of discourse' where the presence of one word usually predicts or triggers the presence of another. Indeed, as indicated by McArthur (1992) at *collocation*, a linguist like Firth used to make a distinction between semantic association only, called 'collocation', and syntactic association (including semantic association, I assume), called 'colligation'. It seems to me that the notion of building block (and also ready-made, for that matter, since it also seems to suggest adjacency) is inappropriate for the collocations in the Firthian sense. Rather, we should more vaguely be speaking of *mutual expectancy facts* that the learner has to learn in order to be fully proficient.

<sup>6</sup> In the 50-million word free BNC sample the total number of occurrences of *shoe shop(s)* is 59, while the total number of *gift shops(s)* is 118.

<sup>7</sup> In the same BCN sample *shop front(s)* scores only 39 hits, whereas *shop window(s)* scores 272. LDOCE4 does mention *shop window*, however, but only as a collocation under *window*!

<sup>8</sup> The BBI, however, unhesitatingly and surprisingly lists solid compounds like *bookshop*, *pawnshop*, *barbershop*, and under *storm* we find e.g. *rainstorm*, *sandstorm*, *snowstorm*. If we call *bookshop* a (= one) 'word', it is obviously not a collocation, for collocations are combinations of words.

### 3 The OCDSE and the MLDs: discussion of the 'listables' or 'ready-mades'

In this section I will compare the noun *storm* as treated in OCDSE and the MLDs. I will (admittedly to some extent intuitively) establish those combinations that are to be considered form-meaning pairings worth being listed as full entries in a dictionary, because they are associated with a 'certain salient type of situation or phenomenon' (see above), that is, because they stand for more or less established fixed and recurrent concepts (hence needing frequent reference) and are therefore more than ordinary collocations. This procedure will involve a comparison of the list of OCDSE collocations with the data in the five MLDs, either offered as separate entries or otherwise.

The OCDSE entry is as follows (minus some example sentences):

#### **OCDSE**

#### **STORM**

##### 1. PERIOD OF BAD WEATHER

- ADJ. bad, big, devastating, disastrous, ferocious, fierce, great, heavy, raging, severe, terrible, tremendous, violent / approaching, gathering, impending ... / freak / autumn, summer, winter / tropical / monsoon / dust, electric/electrical, lightning, magnetic, rain, sand, snow, thunder (also thunderstorm)
- VERB + STORM be in for a ...
- STORM + VERB hit sth, strike sth ... / rage ... / be brewing, be coming / blow up, break, burst / abate, blow itself out / blow over, pass, subside ... / batter sth, buffet sth, lash sth, ravage sth, sweep sth ... / last
- STORM + NOUN cloud / damage, losses
- PREP during/in a/the-
- PHRASES at the height of the storm, the calm/lull before the storm, the eye of the storm

##### 2. violent display of strong feeling

- ADJ. approaching, coming, gathering ... / political
- VERB + STORM arouse, cause, create, provoke, raise, spark, / face / ride out / weather ...
- STORM + VERB blow up, break, burst, erupt / blow over, pass
- PREP. Amid a/the - / - between, - of... / - over

Among the ADJ + STORM combinations the combination *gathering storm* a form-meaning pairing with likely headword status, because of its meaning that is more than its component parts. This is only recognised in MED, which in a subentry under *gathering* remarks 'MAINLY JOURNALISM trouble that is coming' in addition to the analysable meaning 'storm that is coming'. The combinations *dust storm*,<sup>9</sup> *electric(al) storm*, *magnetic storm*, *rain storm*, *sand storm*, *snow storm*, *thunderstorm* should also be considered as having headword status. And in fact this feeling is frequently supported by the MLDs: *dust storm* is found as a separate entry in OALD7, LDOCE4, MED, COBUILD, *electric(al) storm* in OALD7, LDOCE4, CALD, MED, *magnetic storm* in OALD7, *rain storm* in OALD7,<sup>10</sup> LDOCE4,

---

<sup>9</sup> It should be clear that *dust*, *rain* etc. are of course not adjectives proper.

<sup>10</sup> In solid spelling.

CALD, MED, COBUILD, *sand storm* in OALD7,<sup>11</sup> LDOCE4, CALD, MED, COBUILD, *snow storm* in OALD7,<sup>12</sup> LDOCE4, CALD, MED, COBUILD, and *thunderstorm* in OALD7,<sup>13</sup> LDOCE4, CALD, MED, COBUILD.

In addition, these dictionaries list more *ADJ + storm* combinations as separate entries, as in OALD7: *perfect storm* 'an occasion when several bad things happen at the same time, creating a situation that could not be worse'. The conclusion so far can already be that apparently the MLDs tend to consider a number of OCDSE collocations to be not simply collocations but rather real and full-fledged lexical items with headword status, probably due to the fact that they are either semantically not quite predictable (due to e.g. metaphoricisation as in *gathering storm* and *perfect storm*) or else stand for well-known phenomena or concepts (like *sand storm*) that are referred to by means of a conventionalised expression, or a combination of these (as possibly again in *gathering storm*).

The OCDSE combination *political storm* (a possible candidate for headword status, because of its frequency and conventionalised metaphorical sense) is found in none of the MLDs as a separate entry. It is interesting, though, to find that other quite comparable combinations with *political* were in fact deemed to be worthy of headword status, like *political action committee*, *political asylum*, *political correctness*, *political economy*, *political geography*, *political science* (e.g. OALD7), all to be found in OCDSE under the nouns concerned. Again, the dividing line between *word combinations that are not themselves headwords* (cf. the OCDSE definition) and *word combinations that have coalesced into new fixed and stable units that are themselves headwords* (seemingly the reasoning behind the MLDs) is drawn differently in OCDSE and the MLDs.

OCDSE mentions only three combinations with *storm + NOUN*: *storm cloud*, found as a separate entry in OALD7, LDOCE4, CALD, MED, COBUILD, *storm damage* and *storm losses*, found in none. The former (*storm damage*), with 20 occurrences in the 50 million BNC sample, would perhaps deserve headword status, given that it is established as a well-known term for insurance companies. As might be expected, the MLDs mention more of these combinations as separate entries: *storm door*, *storm window* (OALD7), *storm cellar*, *cloud*, *door*, *drain*, *lantern*, *window* (LDOCE4) *storm cloud*, *door*, *window* (CALD), but no further finds in MED or COBUILD.

Other nominal phrases in OCDSE which might be considered for headword status are *the calm/lull before the storm* and *the eye of the storm*. The former is of course explained in all five MLDs, but only under the headwords *calm/lull*, which gives it more the status of an idiom, as explicitly in OALD7 (for *calm before the storm*) though this same OALD7 treats *lull before the storm* only in an example sentence with highlighting bold face – a clear demonstration of the confusion surrounding headword status. *The eye of the storm* is interesting because it has not been afforded headword status in any of the MLDs, though it is found in all

<sup>11</sup> In solid spelling.

<sup>12</sup> In solid spelling.

<sup>13</sup> Idem.

of them under *eye*. The reason for this difference in treatment becomes evident when we look at e.g. OALD7: '*a / the [eye] of a / the storm, tornado, hurricane, etc. a calm area at the centre of a storm, etc*': the phrase is not entirely fixed and admits some variation as long as it refers to a case of violent winds after *eye of*.<sup>14</sup> This somewhat less fixed character makes lexicographers apparently wary of assigning it full entryhood. All this applies of course with even greater force to the combinations with verbs and prepositions, which have no completely fixed quotable form.

In other words: there is a great deal of uncertainty about the difference between collocations and genuine lexical items, with apparently cases like '*a / the eye of a / the storm, tornado, hurricane, etc.* halfway between these two. This makes it worth our while to give some attention to the definition of lexical item or 'headword'.

#### 4 Some concluding remarks about lexical items alias headwords.

As remarked in the first section, command of a language means being able to handle items from a large list of memorised items, including words/lexical items but at least also collocations. Dictionaries have traditionally restricted their strictly alphabetised lists to lexical items (i.e. headwords) but it has never been made quite clear what the criteria are for headword status. This problem for English lexicography applies especially to compounds or compound-like combinations, as already pointed out in Van der Meer 1996. It has long been a tradition in British lexicography (less so in American lexicography) to hide compounds with open spelling under the first headword (if mentioned at all).<sup>15</sup> The evidence above shows this is now changing (also see already e.g. OALD6 and COD9). Yet, the change is perhaps not radical enough.

It is clear, though, that the MLDs – to varying degrees – have gone further along with this trend than OCDSE (and the BBI). What are – and should be then, also in the light of this current practice – the criteria for headword status of combinations of words with open spelling?<sup>16</sup>

I believe the following should be kept in mind: WHEN A COMBINATION (NOT SPREAD OUT OVER TWO OR MORE SYNTACTIC PHRASES)<sup>17</sup> HAS AN UNPREDICTABLE AND/OR SPECIALISED MEANING, AND/OR IS RECURRENT AND/OR REFERS TO FAMILIAR PHENOMENA OR CONCEPTS, IT SHOULD IN PRINCIPLE BE ACCORDED HEADWORD STATUS. Whether such combinations actual-

---

<sup>14</sup> In OCDSE we also find *the eye of the hurricane* under *hurricane*.

<sup>15</sup> This, by the way, is either the consequence or the cause of a slightly different interpretation of the concept of 'word' by English speakers.

<sup>16</sup> I am ignoring here the cases with hyphenated spelling, which are certainly real headwords, though the rules for hyphenation are far from clear.

<sup>17</sup> This is a little theory-dependent, but what I have in mind is combinations like *take a good look at*, which might be called a verb phrase (or predicate) consisting, however, of a verb and its object, the noun phrase. I think it would be wise to exclude them from the wordlist. This is not to exclude fixed combinations like expressions or proverbs from list status; they are, however, to be distinguished from headwords/lexical items, I feel. The criteria for this are to be developed.



ly end up in the wordlist may have to depend on the size of the dictionary and/or its purpose. Thus, even when the aggregate meaning is predictable but the other conditions apply, a combination (compound really) has headword status. Moreover, in many cases, even when one might consider the aggregate semantics predictable and regular, it is often one out of more possibilities, so that after all it might be said that the choice from these alternatives is unpredictable from a purely linguistic point of view. Take for example *hair trigger*: there is nothing from a purely linguistic point of view preventing the non-existent interpretation \*'trigger made of hair', just as *hair shirt* is a shirt made of hair (with additional, linguistically totally unpredictable, semantics involving self-punishment prompted by religious fervour).

A final word is in order about looser combinations like *strong wind(s)*. It is a comparatively frequent combination,<sup>18</sup> which would be a strong qualification, according to the above criteria, for headword status (or perhaps more properly called entry status – either as a subentry or as a main entry), though its semantics is rather predictable and unexceptional. Because of the latter fact, it is understandable that at present the MLDs do not list it in their macrostructures, though for example OALD7 does list *strong-arm* (adj.), *strongbox*, *strong force*, *strongman*, *strong-minded*, *strongroom*, *strong safety* and *strong-willed*. Of these, *strong-minded* and *strong-willed* do not differ too much from *strong winds* in their semantic transparency and hence predictability, but their hyphenated spelling does admittedly make them word-like than *strong winds*. This, of course, leads to the question why they have been hyphenated – surely, one may assume, because they were considered to be 'words', but why? The other items in this list are all semantically specialised and therefore unpredictable, while some (*strong-arm*, *strongbox*, *strongman*) in addition have typical word stress. If we conclude from these examples that apparently the lexicographer's reasoning in the case of *strong winds* was based on semantic transparency, one starts to wonder why then cases like *shoebox*, *shoelace* and *shoemaker* – which, if anything, are at least just as transparent as *strong winds* – did earn full entry status. The answer may well depend on an analysis of the concept of the vague notion of 'word' in English spelling and lexicography. The customary 'open' spelling of many compounds in English seems to suggest a different concept of 'word' from for example Dutch or German. A clearer notion of the English concept of 'word' might make it easier to take decisions about including or excluding items from the macrostructure (or word-list – where 'word' begs the question).

## References

### A. Dictionaries

- BBI = Benson, Morton, Evelyn Benson & Robert Ilson (1997). *The BBI Dictionary of English Word Combinations*. Amsterdam/Philadelphia: John Benjamins.
- CALD = *Cambridge Advanced Learner's Dictionary*; ed. by Patrick Gillard, Kate Woodford & Guy Jackson; Cambridge: Cambridge University Press, 2003 [see CIDE].
- COD: Della Thompson (Ed.) (19959). *The Concise Oxford Dictionary of Current English*. Oxford: Oxford University Press.

<sup>18</sup> It is found 144 times in the free 50 million word sample of the BNC. The singular *strong wind* scores 85.

- COBUILD4 = *Collins Cobuild Advanced Learner's English Dictionary*; Fourth Edition, ed. by Michela Clari. HarperCollins: Glasgow. 2003.
- LDOCE4 = *Longman Dictionary of Contemporary English*; Fourth Edition with Writing Assistant, ed. by Della Summers. Longman: Harlow. 2005
- MED = *MacMillan English Dictionary for Advanced Learners*; ed. by Michael Rundell. Oxford: MacMillan Education Ltd. 2002
- NID3 = *Webster's Third New International Dictionary of the English Language Unabridged*. Ed. by Philip Babcock Gove. Springfield, Massachusetts, USA. 1961/1981
- NODE = *The New Oxford Dictionary of English*; Ed. by Patrick Hanks. Oxford: Clarendon Press. 1998.
- OALD6 = *Oxford Advanced Learner's Dictionary of Current English*. Sixth Edition, ed. by Sally Wehmeier. Oxford: Oxford University Press. 2000.
- OALD7 = *Oxford Advanced Learner's Dictionary of Current English*; Seventh Edition, ed. by Sally Wehmeier. Oxford: Oxford University Press. 2005
- OCDE = Jonathan Crowther, Sheila Dignen, Diana Lea. *Oxford Collocations Dictionary for Students of English*. Oxford: Oxford University Press. 2002
- OED Online = *The Oxford English Dictionary*, 2nd ed.; <http://dictionary.oed.com/entrance.dtl>

#### B. Other references

- Benson, M. (1985), 'Collocations and Idioms', in Ilson, R. (ed.) *Dictionaries. Lexicography and Language Learning*. Oxford and New York: Pergamon Press [61-8].
- Jackson, H., Zé Amvela, E. (2000), *Words, Meaning and Vocabulary. An Introduction to Modern English Lexicology*. London and New York, Cassel.
- McArthur, T. (ed.) (1992), *The Oxford Companion of the English Language*. Oxford, Oxford University Press.
- Meer, G. van der (1996), 'How alphabetical should a dictionary be? (the case of HIGH and its combinations in some dictionaries)'. *SYMPOSIUM ON LEXICOGRAPHY VII. Proceedings of the Seventh Symposium on Lexicography May 5-6, 1994 at the University of Copenhagen (Lexicographica, Series Major 76)*, edited by Arne Zettersten and Viggo Hjornager Pedersen, pp. 183-197.
- Meer, G. van der (1998), 'Collocations as one particular type of conventional word combinations. Their definition and character'. *EURALEX '98 PROCEEDINGS I-II, Papers submitted to the Eighth EURALEX International Congress on Lexicography in Liège, Belgium. Voll*, pp. 313-322.
- Klotz, M. (2003), [review of] *Oxford Collocations Dictionary for Students of English*', in *International Journal of English* 16, pp. 57-61.
- Steele, J. (ed.) (1990), *Meaning-Text Theory. Linguistics, Lexicography, and Implications*. University of Ottawa Press, Ottawa etc.
- Warren, B. (2004), 'A Model of Idiomaticity', [www.hum.au.dk/engelsk/naes2004/download\\_paper.html?ID=1](http://www.hum.au.dk/engelsk/naes2004/download_paper.html?ID=1)