An Analysis of the *Collins COBUILD English Dictionary,* New Edition[1]

Hideo Masuda   Makoto Kozaki
Naoyuki Takagi   Kyohei Nakamoto
Rumi Takahashi

1. Introduction

When *Collins COBUILD English Language Dictionary* (abbreviated to *COBUILD*[1]) was issued in 1987, users in Japan (most of them were school teachers and lexicographers) were surprised because they found it very different from the conventional dictionaries. What surprised them most was the way *COBUILD*[1] listed the entry words. It employed the one word, one entry policy. That policy was applied so rigidly that no distinction was made between homonyms and homographs. Thus in the main entry word **seal**, two unrelated words that have different etymologies were listed: **1** A **seal** is **1.1** a special design that is the official mark of a person or organization. **4** A **seal** is also a large animal that eats fish and lives partly

on land and partly in the sea, usually in cold parts of the world. Similarly, two words with different pronunciation but the same spelling entered under the same headword. **Lead** /liːd/ and /lɛd/, and **pace** /peɪs/ and /peɪsi¹/ were the case.

COBUILD[1] was criticized for these reasons. At the same time it was praised for its innovation in giving definitions in full sentences together with the 'authentic' illustrative examples based on a large English corpus. It is also praised, and criticized at the same time, for its innovative transcription of reduced vowels, and nasals. Numerical superscripts such as /ə¹/~/ə⁹/, /ɪ¹/~/ɪ⁵/, /m¹/, /n¹/, etc. might be a good way to show the shifts in pronunciation of weak vowels and nasals, but they are not userfriendly. These figures impose unnecessary burdens on the learners. Underlining to indicate a shift in stress was easy for learners to understand, was still comprehensive in its application. It deserved high praise.

Grammatical information, such as verb patterns given in 'Extra Column,' was fairly detailed and informative. However, there were cases where verb patterns and illustrative examples did not correspond. The code 'V-ERG' (ergative) has been introduced for the first time into the dictionary. Since there are many verbs that are used either as intransitives and as transitives, this new code is very useful for giving grammatical information of that kind.

On the whole, critics in Japan have been favorable to *COBUILD*[1] in spite of the shortcomings, many of which all first edition dictionaries share. The dictionary has been welcomed by English teachers in Japan, but we wonder how many learners appreciate it. The target users of the dictionary seem to be advanced learners rather than beginners.

We believe the Bank of English, which is growing at a rate of around 5 million words per month (according to Internet information), made it possible to refine the definitions supported by totally renewed illustrative examples. We are very interested to know what efforts have been made to eliminate the previous shortcomings, and what further innovations have been made in the New Edition, and what new features have been added for a better understanding of the English language.

2. Entries and Information in the Extra Column

This section deals with the entries in the 2nd edition of *Collins COBUILD English Dictionary*, first from the quantitative point of view. In addition to the quantitative analysis of the entries, the qualitative aspect of the 2nd edition is also taken into account, that is, the way the entries are presented. This section also looks at the information given in the Extra Column, especially at the introduction of *Frequency Bands*.

2.1. Entries

The total number of pages in this dictionary has increased by 248 pages from the 1703 of the first edition to 1951 in the present volume. Although the number of lines per page decreased from 174 to 164, this increase in page numbers results in an increase of about 236,000 lines in total. Needless to say, this expanded space means not only more detailed explanations of words and phrases already in the dictionary but also the introduction of a number of new entries.[1]

2.1.1. The total number of main entries in $COBUILD^2$ is not exactly shown except that it claims on the backcover that it covers "over 75,000 references". If we naïvely regard the expression of "references" as "entries," this figure means 5,000 additional entries to the first edition. However, as it was pointed out in Kojima et al. (1989: 46), $COBUILD$'s way of counting entries is somewhat problematic. In order to check the validity of its claim of more than 75,000 references, we used a sampling test.

2.1.2. First, sample pages were taken at every 40 pages of $COBUILD^2$, and they amounted to 48 pages. Then it was found that the average number of main entries was 17.3 words per page. In order to figure out the total number of main entries in $COBUILD^2$, this average word number (17.3) was multiplied by the total number of pages in $COBUILD^2$, which is 1951. The answer thus obtained was 33,752, which is very far from the number indicated above. Even if the total number of run-on entries (4,682;

---

1) $COBUILD^2$ (p. ix) says that 'COBUILD made available a lot of space for new and additional entries, by increasing the size of the book and also by more efficient presentation'

2.41 per page) and that of sub-entries (2,556; 1.31 per page) are added together, the result is 40,990, which pales in comparison with the dictionary's claim of 75,000 references. This result reveals that $COBUILD^2$ is not exceptional in taking advantage of the now-prevalent way of counting dictionary entries: namely, regarding all the information in bold face as entries[1].

2.2. New entries and deleted words

$COBUILD^2$ employs a computer corpus of over 200 million English words, which is roughly ten times larger than the one used in the first edition. This expansion of the size of the corpus, it is claimed, is reflected in the choice of entries.

2.2.1. First, let us look at the main entries included in $COBUILD^2$ but not in $COBUILD^1$. Newly adopted main entries amount to 101 in the sample above used. Only 2 words of these 101 new entries are treated as run-on entries in the first edition. And another 2 words are described as part of the definition of their main entries in $COBUILD^1$. As a result, entirely new entries account for about 96% of the total. On the other hand, 76 main entries in $COBUILD^1$ have disappeared, with 10 of them treated as run-ons and 5 described as part of examples in $COBUILD^2$.

2.2.2. Now we will get down on a closer look at the new entries. $COBUILD^2$ (p. ix) says that it "is <u>not</u> a historical record of the language, and it is not a list of all the peculiar words" such as those used in a crossword (emphasis in original). The first point to be noted concerning the new main entries is that there are a lot of compounds among them. About half of the main entries newly adopted in $COBUILD^2$ are compounds. For example, there are 8 new compound entries at page 40 of $COBUILD^2$: air-drop; airfare; airframe; air freshener; airhead; airplay; air power; airshow. This fact may suggest that $COBUILD^2$ is quite willing to treat multiple word expressions as main entries.

2.2.3. As is pointed out in Kojima et al. (1989: 64), the entries in $COBUILD^1$ are quite neutral in terms of regional, technical and stylistic

---

1) See Landau (1984: 84).

varieties. This means that it tends not to include words used by a particu-
lar group of people or in a particular social context. In line with this
tendency, *COBUILD²* (p. ix) claims that it 'gives priority to the English of
most general utility worldwide.' As far as new entries are concerned, how-
ever, *COBUILD²* seems to put an emphasis on American English because
8 of the new main entries in our sample are labeled as "American" while
only 4 of them are labeled as "British."[1] This might be a reflection of this
dictionary's own claim that it covers a lot of American usage as well as
British English.

Another noticeable fact about the new main entries in *COBUILD²* is
that it seems to be in favor of informal words rather than formal ones. In
our sample of 101 new main entries, 8 entries are labeled as "informal" or
"very informal" (*e.g. blag, chutzpah, ciao, ciggy, hiya, knacker, pic,* and
supremo), whether they are British English or American, while only 2 are
marked as "formal." (*e.g. curatorial* and *proportionality*) This might be the
result of improved treatment of the spoken language in the corpus.

2.3.   Changes in Entry Organization

The way of arranging and presenting entries is basically the same as that
of the first edition. That is to say, main entries are presented in bold face,
one letter protruding to the left side margin. And the inflected forms of the
main entries follow in smaller bold face. The same symbol ● which is
employed in the first edition is still used in *COBUILD²* to introduce set
phrases and cross-references.

2.3.1.   There are slight changes in the use of symbols. Derived words are
treated as run-on entries following the symbol ◆, which has replaced ◇ of
the first edition. Another change is the adoption of the symbol ▶, which is
used in three different ways. First, it represents a change in word class
involving no change in meaning. Second, it introduces a meaning closely
connected with another meaning. And third, it introduces phrasal verbs
which have the same meaning as the headword verb. However, phrasal
verbs are generally given the status of sub-entries and presented when the

1)   *LDCE³* and *OALD⁵* are also used to examine regional varieties and usage labels.

entire description of the main entry ends, just as in the first edition.

2.3.2.   The most noticeable change in the presentation of main entries of
COBUILD² is in the introduction of *Superheadwords*. Although the first
edition of this dictionary was epoch-making in that it kept to the strict
policy of 'one word, one entry', it raised a question of whether words of
different classes or of different origins should be placed under the same
headwords. *COBUILD²* takes into account the disadvantage pointed out
in reviews of the first edition, and tries to make it easier for learners to find
a word with several meanings. One way which *COBUILD²* has adopted in
trying to solve word-finding problems is through the introduction of
Superheadwords.

Superheadwords is a way of dividing a main entry of different word
classes or with several quite different meanings into sub-entries. Each
superheadword is marked with a number and followed by a brief explana-
tion of a given word. For example, *order* is divided into three sections:
"**order 1** subordinating conjunction uses"; "**order 2** commands and re-
quests": and "**order 3** arrangements, situation, and groupings". As a re-
sult, *COBUILD²* looks, in appearance, more like a conventional dictionary
in that there exist several headwords for the identical form, just as homo-
graphs are marked with superscripts and treated separately in traditional
dictionaries.

2.3.3.   The introduction of *Superheadwords* seems quite innovative at first
glance, but whether or not this apparently novel way of presenting
polysemous items will help ease word-finding problems is a different mat-
ter. The first question to be raised concerning the use of *Superheadwords* is
that there seems to be no explicit principle which decides whether or not a
superheadword is employed for a given item with multiple meanings or of
distinct origins. For example, *light* is divided into three sub-entries such as
"**light 1** brightness or illumination", "**light 2** not great in weight, amount,
or intensity", and "**light 3** unimportant or not serious". The differentia-
tion between **light 1** and **light 2** is quite reasonable, even to the non-native
speaker's intuition, but the distinction between **light 2** and **light 3** is
unpersuasive. The meaning "unimportant or not serious" of **light 3** is
rather easy to guess from the meaning of **light 2**, especially from the 11th

paragraph[1] through the process of metaphorical semantic extension.[2] There seems no positive reason to put them into different sections. In contrast to this, *heavy* is treated under a single headword without using a superheadword. If there is any need to separate **light 2** and **light 3**, then why is the 17th paragraph of **heavy** not treated as a different sub-entry?

There is another question about *Superheadwords*. Despite the adoption of this new way of presenting headwords, different word classes are put under the same headword, as in *COBUILD*[1]. For example, **light 1** contains verb and adjective uses as well as the noun ones. In contrast to this, superheadwords are sometimes employed according to the differences in word class. *Back* is divided into three parts as: "**back 1** adverb uses"; "**back 2** opposite of front; noun and adjective uses"; and "**back 3** verb uses". Such a fluctuating policy of employing superheadwords as this will be quite confusing for the users of this dictionary.

2.3.4. As to the organization of run-on entries, the problem pointed out in Kojima et al. (1989: 49) still remains unsolved. In this dictionary, derived words appear sometimes as main entries and sometimes as run-on entries. The adverb *positively* appears 4 times as a run-on entry while, at the same time, it appears independently as a main entry. Furthermore, in the 8th paragraph of the main entry *positive*, a cross-reference to the adverb *positively* is offered, but that is a little confusing because it is not immediately clear which *positively* is referred to.

Similar word-finding problems may occur when we look up verb-particle combinations such as *put on* and *take off*, because of this dictionary's treatment of run-on entries and sub-entries, i.e. phrasal verbs.[3] For example, we cannot find any obvious reason why the example *Take off the price of the house, that's another five thousand*, which is in the 10th paragraph of the verb *take* is treated as an independent use of the headword verb and not as a use of the phrasal verb *take off*.

1) This definition reads as follows: "**11** If you describe the result of an action or a punishment as **light**, you mean that it is less serious or severe than you expected."
2) As to the metaphorical semantic extension of *light* and *heavy*, see Kozaki (1997).
3) As has been mentioned in 2.3.1, verb-particle combinations are treated as run-on entries or sub-entries.

2.4. Information in the Extra Column

It can be said that one of the most characteristic features of the *COBUILD*[1] is its efficient use of the Extra Column. *COBUILD*[2] also makes use of this informative space, but with a slight change. As grammatical information and Pragmatics are examined in detail in other chapters, we focus on the presentation of related words in main entries and frequency of word uses.

2.4.1. In the Extra Column of *COBUILD*[1], the symbols ⇑, = and ≠ are used to introduce respectively a superordinate word, a synonym, and an antonym of a given word. *COBUILD*[2], however, retains only the latter two of those symbols. Superordinates are no longer offered in *COBUILD*[2], which is quite deplorable. Although the presentation of superordinates in the first edition was severely criticized as unsystematic and far from satisfactory[1], it cannot be denied that this innovation is quite helpful for learners as well as inspiring to lexicographers.

2.4.2. One of the most noticeable innovations in *COBUILD*[2], along with the introduction of *Superheadwords*, is the adoption of *Frequency Bands*. This is an index system of the frequency of word uses in the Extra Column by means of five diamond-shaped symbols. The most frequently used 700 words are marked with all five diamonds in black, and 1,200 words of a little less frequency are given four black diamonds and a white diamond. Three black and two white diamonds are given to the 1,500 words at the next frequency level. 3,200 words used still less frequently are marked

Frequency Bands	Total Number of Words	Examples
◆◆◆◆◆	700	the, and, of, to, like, go, paper, return
◆◆◆◆◇	1,200	argue, bridge, danger, female, obvious, sea
◆◆◆◇◇	1,500	aggressive, medicine, tactic
◆◆◇◇◇	3,200	accuracy, duration, miserable, puzzle, rope
◆◇◇◇◇	8,100	abundant, crossroads, fearless, missionary

1) See Kojima et al. (1989: 107)

with two black and three white diamonds. And, finally, a combination of one black diamond with four white marks 8,100 words of still lower frequency.

These five levels of frequency, combined with no markers for least-frequently-used words, make it possible to differentiate 6 levels of frequency of words in *COBUILD²*. No doubt, they are quite useful for learners, because they tell them which word should be learned first.

2.4.3. Although there is no denying the fact that information given by *Frequency Bands* is quite helpful for learners of English, there remain some points to be improved. The first point is that it only offers the frequency of uses of a headword, not of separate meanings of the word. For example, the headword *effect* is marked with five diamonds, indicating that it is among the most frequently used words of English, but when we look at the fifth paragraph of this word, we find that the verb use of it is presented there. This raises a question of whether the verb *effect* is so frequently used as to be included among the most important 700 words. This way of presenting word frequency is quite confusing in that there is a possibility that learners at the beginner level will regard all the different meanings of this headword as most frequently used. The cause of this problem lies partly in the way this dictionary organizes entries. As has been mentioned in 2.3., *COBUILD²* basically employs "one word, one entry" policy. Because of this policy, the noun uses of the word *effect* and the verb use of it are put together under the same headword.

(M. Kozaki)

3. Pronunciation
3.1. Overview

The new edition has a strong prescriptive orientation, which is stated in its guide to pronunciation on p. xxxviii (this section is henceforth referred to as the Guide) as follows:

> Our aim has been to provide a pronunciation key that is accurate, clear, and simple. The basic principle underlying the suggested pronunciations is 'If you pronounce it like this, most people will understand you'.

Pronunciations in the revised edition are provided using the phonetic symbols adapted from those of the International Phonetic Alphabet (IPA), following the convention of the English Pronouncing Dictionary (EPD, 14th Edition revised by A.C. Gimson and S.M. Ramsaran)[1].

Unlike the previous edition, which described only Received Pronunciation (RP), the current edition has both RP and General American (GenAm) in its descriptive scope. Systematic vowel differences between the two major accents of English are explained in the Guide and left unmarked throughout the main body of the dictionary. Only when it is necessary, a separate transcription for General American is given after the code AM. Although it is stated that "Where more than one pronunciation is common in British English, alternative pronunciations are also given", this is done only sparingly. In consequence, the present edition provides only one pronunciation key for the majority of words. This is a reasonable choice given the above mentioned editorial principle.

As for the description of GenAm, there seems to be some room for improvement. For most of the so called "CLOTH-words" (Wells, 1982), the American pronunciation with vowel /ɔː/ is shown after the code AM. However, for the word *fog*, this American variation is not shown. The entries for *issue* and *ate* are /ɪsjuː, ɪʃuː/ and /et, eɪt/, respectively, indicating erroneously that /ɪsjuː/ and /et/ are possible American pronunciations.

3.2. Use of superscript r

In this dictionary, the superscript /ʳ/ shows that "i) in RP /r/ is pronounced only when it is followed by a vowel; ii) in GenAm, r is always pronounced." The Guide provides further examples such as *fire, flour, fair, near,* and *lure* and states, for example, that *fire* is pronounced /faɪə/ in

---

1) Cobuild uses the symbol /oʊ/ for the vowel in *go* and *hope*, and the rationale behind this, stated in the Guide, is that this symbol is the hybrid of RP /əʊ/ and General American /o/. This choice, however, fails to give an accurate transcription for the RP pronunciation. Since phonetic symbols for RP is used whenever there is a systematic sound difference between RP and GenAm elsewhere (e.g. RP /ɒ/ and GenAm /ɑ/ for *hot*), users might believe that /oʊ/ is the RP pronunciation. This can be avoided by the use of /əʊ/ instead of /oʊ/.

RP and /faɪr/ in GenAm. This is done to avoid the confusion that /r/ is pronounced after /ə/ in GenAm since these examples contain a diphthong or a triphthong that end with /ə/ in RP. The superscript /ʳ/ is also used to represent pronunciations of words such as *hard* /hɑːʳd/ and door /dɔːʳ/.

The introduction of the superscript /ʳ/ makes possible a concise description of both RP linking /r/ and GenAm pronunciations, but it has certain shortcomings. When it comes to describing the monophthong that appears in *bird, hurt*, etc., it fails to accurately represent the phonetic difference between RP and GenAm. The vowel in *bird* in GenAm is a monophthong, and as Takebayashi has pointed out (Takebayashi, 1984; Higashi et al. 1986), the use of symbol /ɜːʳ/ is extremely misleading since this gives the impression that the vowel is a diphthong pronounced /ɜːr/ in GenAm.

This notation is also problematic in transcribing words in which this vowel occurs before another vowel as in *stirrer*. This word has a single pronunciation key in the dictionary: /stɜːrəʳ/. Here, the /r/ after the vowel /ɜː/ is no longer in superscript, and no RP-GenAm difference is obvious from the transcription as far as /ɜː/ is concerned. Moreover, when /əʳ/ is interpreted as /r/ for GenAm, we obtain /stɜːrr/, which suggests that this word is monosyllabic with two /r/'s at the end.

One solution is to use an additional symbol such as /ɚ/ or /ɝ/ for the GenAm r-colored vowel as is suggested by Takebayashi (Takebayashi, 1984; Higashi et al. 1986) and used in several English-Japanese dictionaries edited by him (e.g. Lighthouse English-Japanese Dictionary, New College English-Japanese Dictionary). This approach makes explicit the differences between RP and GenAm: /faɪə/ vs. /faɪɚ/; /stɜː/ vs. /stɚː/; /stɜːrə/ vs. /stɚːrɚ/. (For further discussion on practical merits pertaining to the use of /ɚ/, see Takebayashi, 1984)

3.3. Reduced Vowels

Unlike the previous edition which used a complicated system of superscript numerals, the present edition cleverly utilizes italicization to represent various patterns of vowel reduction. To show that the first vowel in "accept" can vary from /æ/ to /ə/, symbol /æ/ is italicized as in /æksept/. Not only full vowels such as /oʊ/ /eɪ/, and /uː/ as in *notorious, candidate,*

and *fortune*, but also reduced vowels /ɪ/ and /ʊ/ are sometimes italicized, indicating that reduction can go toward /ə/. This convention saves space and provides an accurate description of vowel reduction at the same time and we find it quite suitable for a dictionary compiled for nonnative learners of English.

3.4. Weak forms

The present edition gives comprehensive descriptions of frequent weak forms across various parts of speech such as articles, personal pronouns, auxiliary verbs, prepositions, and conjunctions. It first gives a weak form, and then adds its strong form after "STRONG" as in /ənd, STRONG ænd/.

Weak forms given in the new edition, however, seem to be limited to frequent ones. For example, weak forms of "your" and "their", both of which are marked as "occasional weak forms" in EPD, are not given. Personal pronouns that begin with /h/ sound (e.g. "him" and "her") often have weak forms whose initial /h/ is dropped, but this possible /h/ deletion is not mentioned. This, again, seems to stem from the prescriptive orientation of the present edition.

3.5. Stress

The present edition follows the unique system of stress description of the previous one, with minor modifications. For polysyllabic words in which primary stress comes after secondary one(s) as in *disappointing* and *disqualification*, all stressed syllables are underlined as in /dɪsəpɔɪntɪŋ/ and /dɪskwɒlɪfɪkeɪʃᵊn/. The fact that the final underlined vowel receives primary stress in citation forms is stated in the Guide and this is regarded as understood in the main body of the dictionary. In the previous edition, stressed vowels were printed in bold and underlined, but this redundancy, which was pointed out by Takebayashi (Kojima, et al. 1989) has gone in the new edition[1].

As Takebayashi pointed out (Kojima, et al. 1989), this system has the

---

1) Underlining is also used to indicate stress patterns for compounds such as *Christmas cake, Christmas Eve*. This was not done in the previous edition.

advantage of indicating possible stress shift accurately. Indeed the Guide has a concise and excellent explanation of how stress shifts depending on contexts in which a given word appears. As two important features of underlined syllables, the Guide mentions the following:

1)   They can take primary or secondary stress in a way that is not shared by the other syllables

2)   Whether they are stressed or not, the vowel must be pronounced distinctly; it cannot be weakened to /ə/, /ɪ/, /ʊ/.

Since these features are shared by most monosyllabic words (i.e. except the ones that have weak forms), vowels in them are also underlined.

One unfortunate consequence of this convention is that secondary stressed vowels that come after primary stressed ones in citation forms cannot be underlined as in "ridicule" /rɪd/kjuːl/, in spite of the fact that the final /uː/ must be pronounced distinctly and it does take secondary stress. Further, when such secondary stressed vowels are either /ɪ/ or /ʊ/ as in *cataclysm* or *outlook*, one cannot tell whether they are stressed or weak vowels.

(N. Takagi)

4.   Definitions

4.1.   Sentence Definitions

4.1.1.   Minor Changes in Sentence Definitions   "The most distinctive feature of the original dictionary was the use of full English sentences in the definitions", said the editor in chief proudly.[1] It is true, though sentence definitions are no longer the trademark of *COBUILD* when its rivals are so keen on following suit. In fact, he has made "no apology for full sentence definitions".[2] On the contrary, they appear to have been further improved in the current edition. A few of the (rather minor) changes help the user read definitions more easily.

1)   *COBUILD*[2], p. xi.
2)   Ibid.

4.1.1.1.   In the 1987 edition, the first part of the definition was not repeated if a word had several senses in a particular pattern of use.[1] In the 1995 edition, however, all definitions start on a new line with a whole definition number, making a separate paragraph each other:

magazine
    1st ed.   A **magazine** is **1** a publication . . . **2** a topical news . . . **3** a compartment in a gun . . . **4** a building . . .
    2nd ed.   **1** A **magazine** is a publication ...
            2 On radio or television, a **magazine** or **magazine programme** is a programme . . .
            3 In a gun, the **magazine** is the compartment . . .
            4 A **magazine** is a building . . .

Similarly, 'sub-paragraphs' introduced by numbers with a decimal point have been abolished:

nod
    1st ed.   If you **nod, 1.1** you move your head . . . **1.2** you bend your head . . . **1.3** you bend your head . . . **1.4** . . .
    2nd ed.   **1** If you **nod**, you move your head . . .
            2 If you **nod** . . . , you bend your head . . .
            3 If you **nod**, you bend you head . . .
            4 . . .

For separation and ordering of different senses, see 4.2.

4.1.1.2.   In the first edition, definitions were sometimes interrupted by phonetic descriptions. This never happens in the new edition:

abatement
    1st ed.   **abatement** /əbeɪtmə[2]nt/ means . . .
    2nd ed.   **abatement** /əbeɪtmənt/. **Abatement** means . . .

4.1.1.3.   In the former edition, the adverb *also* was often used to connect different paragraphs under the same headword. For instance, the second sense of **immediately**, "You use **immediately** . . .", was followed by the third, "You also use **immediately** . . .". The "also" presupposed that the user would read all definitions consecutively. The new edition seems to

1)   *COBUILD*[1], p. viii.

have avoided the paragraph connector, separating paragraphs from each other.[1]

We welcome the new layout, which will make it easier for the user to seek the information he wants and, of course, much easier for him to read.

Incidentally, it should be mentioned that different forms of a compound are now fully spelled out. Compare:

curtain call
> 1st ed.   also spelled with a hyphen.
> 2nd ed.   also spelled **curtain-call.**

The new edition is more 'user-friendly' in this respect.

4.1.2.   More Important Changes The new features mentioned so far all concern manners of presentation — in other words, the 'craft' of lexicography. The new edition's definitions will now be reviewed more closely.

4.1.2.1.   Some definitions were re-written so that they match the grammatical descriptions provided in the 'Extra Column'. First of all, some common nouns are defined with a plural noun head[2]:

cockle
> 1st ed.   **A cockle** is an edible shellfish.   N COUNT: USU PL
> 2nd ed.   **Cockles** are small edible shellfish.   N-COUNT: usu pl

Other common nouns are defined in a similar way, even though they do not have a 'usu pl' label:

dragonfly
> 1st ed.   A **dragonfly** is a brightly-coloured insect . . . N COUNT
> 2nd ed.   **Dragonflies** are brightly-coloured insects . . . N-COUNT

This is because "there are much more examples in our corpus where *dragonfly* is used in the plural form".[3]

Rather unfortunately, however, the new edition is full of inconsistencies

---

1) But *also* is left at **however** (def. 5).
2) Plural nouns like *police* are defined with a plural head both in the first and the second editions.
3) Ms Gwyneth Fox, editorial director, personal communication.

in this respect. *Butterfly* is defined in its singular form; *cod* (N-VAR) as plural but *carp* (N-VAR) and *salmon* (N-COUNT) as singular. The third sense of *flower* has a plural head, whereas the second has a singular (both with the same grammatical label, 'N-COUNT: usu pl').

It would be fairly easy to make the dictionary more consistent: if a particular noun is typically used in its plural (or singular) form, write its definition with a plural (or singular) noun head with an appropriate grammatical label such as 'usu pl' (or 'usu sing'). However, we still have to wonder why *dragonfly* and *cod* are typically plural when *butterfly*, *carp* and *salmon* are typically singular. Is 'The Bank of English' really reliable in this respect?

There are similar inconsistencies in the second edition. Compare the next two excerpts (both from the second edition):

cosmos
> **The cosmos** is the universe; . . . N-SING: *the* N = universe
universe
> The **universe** is . . .   N-COUNT: usu *the* N in sing

The different typefaces may reflect the differences in the grammatical labels supplied (that is, with or without "usu"). However, the following example breaks this rule:

mother country (def. 2)
> If you refer to **the mother country** . . .   N-SING: usu *the* N

It would cause little trouble for the user whether the definite article is printed in bold typeface or not. The real problem lies in the fact that he never knows how different '*the* N' is from 'usu *the* N' or even 'oft[en] *the* N' (e.g. **jitters** "N-PLURAL: oft *the* N").

We must hasten to add, however, that many of the definitions of the second edition match their grammatical labels. Here are some examples:

persistence
> 1st ed.   (def. 1) The **persistence** of something   N UNCOUNT
> is . . .
> 2nd ed.   (def. 2) The **persistence** of something,   N-UNCOUNT:
> . . . , is . . .   usu *the* N *of* n

abdomen

    1st ed.   The **abdomen** is . . .   N COUNT

    2nd ed. Your **abdomen** is . . . N-COUNT: oft poss N

lance

    1st ed.   If you **lance** a boil on someone's body, . . . V + O

    2nd ed. If a boil on someone's body **is lanced**, . . . VB: usu passive

bay[1)]

    2nd ed. (def. 6) If . . . a number of people **are**   VB: usu cont
           baying . . .

These are all good definitions. The next has also been improved, but it still has a problem:

entwine

    1st ed.   If you **entwine** something in   V-ERG: USU PASS,
           something else, . . .   IF + PREP THEN *in/with*

    2nd ed. (def. 1) If one thing **is entwined** with   V-RECIP-ERG
           another thing or if you **entwine** two things, . . .

The second edition's definition, which is obviously better than that in the first, indicates only two of the four possible syntactic behaviours of this verb.

4.1.2.2.   Many sentence definitions still use verbs such as *mean* and *use* to describe meanings, as the first edition did. The use of these verbs not only spoils the value of the defining sentence but could be confusing. Here is an improved definition:

envelop

    1st ed.   To **envelop** something means to cover, surround, or enclose
           it completely.

    2nd ed. If one thing **envelops** another, it covers or surrounds it
           completely.

But **ruin** (def. 1), for example, is defined as "To **ruin** something means to severely harm, damage, or spoil it.", which is virtually the same as the first edition's definition to this sense.

---

1) This sense was not included in the first edition.

Adjectives and adverbs are often defined by means of *mean* (e.g. **early**), while *use* is most often used when explaining the uses of function words (e.g. **in**). The following definitions are confusing for the user — at least for Japanese learners of English:

magic (2nd ed.)

    1 **Magic** is the power to use supernatural forces . . .

    2 You can use **magic** when you are referring to . . .

    3 You use **magic** to describe something . . .

The first definition is a little confusing and the latter two are much more confusing, because *use magic* is an acceptable, though not typical, collocation.[1)] This collocation is especially confusing for Japanese learners of English; there is a strong Japanese collocation, *mahou o tsukau,* whose literal translation is 'to use magic'. The Japanese learner could fail to understand that *use* in the latter two definitions is an explanatory verb.

It would not be very difficult to avoid the use of *mean*,[2)] but it is a troublesome issue how lexicographers might avoid confusion by using a heavy-duty verb *use*.

4.1.2.3.   The discussion in the preceding section makes us realise that there exist 'hidden' problems that are unavoidably created by sentence definitions.

Defining sentences are probably most useful for describing (1) selectional restrictions of a headword:

beautiful

    1st ed.   (def. 1) Someone or something that is **beautiful** is very good
           and pleasing to look at.

    2nd ed. (def. 1) A **beautiful** woman or child is very attractive to look
           at.

or (2) typical collocations:

hay fever

    1st ed.   **Hay fever** is inflammation . . .

---

1) The first sense of **magic** (adj.) in *OALD*[5] reads, "used in or using magic".

2) The above definition could have been "If someone **ruins** something, they severely harm, damage, or spoil it."

2nd ed. If someone suffers from **hay fever**, . . .

or even (3) typical 'users' of the headword:

wild child[1]

> In British English, journalists sometimes use **wild child** to refer a young teenage girl . . .

For the first two headwords, the second edition gives more information than the first.

However, there are problems as far as (2) is concerned. First of all, there are unfortunate inconsistencies. Some diseases are given typical collocates, but others are not. The new edition tells the learner that people may "have a cold, headache, or stomach-ache", but "suffer from amnesia, jet lag, or rhinitis". Yet he never knows whether people may "have a backache, heartache, or toothache", nor does he know whether they may "suffer from AIDS, cholera, or leukaemia".[2]

It will not be a hard task to 'cure' *COBUILD* of such inconsistencies in the next edition. The crucial problem is: When do definitions of words (or compounds) and those of phrases (and even examples) part from each other? Compare the full definition for **hay fever** with that in *OALD*[5]:

> *COBUILD*[2]: If someone suffers from **hay fever**, their nose, throat, and eyes become inflamed, usually because they are allergic to the pollen of some grasses or flowers.
> *OALD*[5]:   an illness affecting the nose, eyes and throat, caused by POLLEN breathed in from the air

It is good that *COBUILD* gives more (encyclopaedic) information than *OALD*, not to mention the important collocation *suffer from hay fever*. However, *COBUILD* does NOT define this compound; it only explains the meaning of the whole expression "someone suffers from hay fever". Unlike *OALD*, *COBUILD* does not tell the user that hay fever is "an illness", which is the genus proximum of this definition (Zgusta 1971:

---

1) The definition is from the second edition. The first edition did not include this compound.

2) Among these only **heartache** has an example which shows a collocate *suffer*.

252). It has also ignored a traditional defining practice, obviously on purpose, that "the definition should reflect the part of speech of the word defined" (Landau 1984: 134).

Defining sentences of this type could allow the lexicographer to write extreme 'definitions' almost to the point of absurdity. Can the following (fake) definition of *taxi*, but not of *take a taxi*, be regarded as 'definition' at all?

> When you take a **taxi**, you go somewhere by it.

This is an over-exaggerated example, but *COBUILD* has already taken a further step. Compare the following definitions both taken from the second edition:

> If you achieve **fame**, you become very well-known. (s.v. **fame**)
> If you **lose** your **temper**, you become so angry . . .
>
>   (s.v. **temper**, def. 5)

To "lose one's temper" is regarded as a (fixed) phrase, both *lose* and *temper* being printed in bold face, whereas to "achieve fame" is not. Clearly, the two explanations equally define the meaning of the 'verb plus noun' phrase, not the entry noun, *fame* and *temper*, respectively. We wonder why *COBUILD* did not treat "achieve fame" as a phrase (see also 6.2).

4.1.2.4. The COBUILD dictionaries use in the definition such pronouns as *you* and *someone* carefully according to the meaning of the entry word. This is also true for the new edition. Some definitions were re-written to improve them. Here is an example[1]:

murder

> 1st ed. (def. 2) If you **murder** someone, you kill them deliberately and in an unlawful way.
> 2nd ed. (def. 2) To **murder** someone means to commit the crime of killing them deliberately.

"If you **murder** someone" is in fact an awkward expression. The awkwardness was avoided in the latter definition, which does not tell the user, however, if it is possible to say in a normal context, "A hungry crocodile

---

1) Other examples are **diddle**, **hostile**, and **send . . . to Coventry** (s.v. **Coventry**).

murdered the poor man". The learner has to make up for the missing information himself; "to commit the crime of killing" is a clue to the answer. "If someone **murders** another person, they kill them deliberately . . . " could have been an alternative explanation.[1]

The alternative definition above reminds us of problems in the use of such personal pronouns as *they, their,* and *them* with reference to a single antecedent. This was what made many reviewers, particularly language teachers, unhappy. Among them Piotrowski (1988: 254) pointed out that such uses could be confusing, as in the above (fake) definition. In the next example the controversial *they* has been avoided:

pretender
> 1st ed.   A **pretender** is someone who claims the right to a particular position which they do not have, when their claim is disputed by other people.
> 2nd ed.   A **pretender** to a position is someone who claims the right to that position, and whose claim is disputed by others.

The first edition also used *he or she,* as in the following extract, which is obviously an awkward definition:

parole
> 1st ed.   (def. 1) If a prisoner is given **parole**, he or she is freed before their prison sentence is due to end, on condition that he or she behaves well.

This has been re-written wisely with a plural noun head[2]:

> 2nd ed.   (def. 1) When prisoners are given **parole**, they are released before their prison sentence is due to end, on condition that they behave well.

It seems that the COBUILD team were careful enough to review their definitions with respect to the uses of pronouns.

Incidentally, the informal use of *you* has become common; other EFL dictionaries have followed suit[3].

---

1) The latter is not a very good sentence definition where an explanatory verb *mean* is used. See 4.1.2.2.

2) Also note that "If" has been replaced with "When".

3) See **inviting** in *LDCE*[3] and *CIDE*, for example.

4.1.2.5.   Generally speaking, definitions in the new edition are simpler and easier to read than those in the first edition. However, simple definitions are not necessarily better definitions. We have found some 'too simple' definitions in the current edition.

It is desirable to avoid superfluous words and phrases:

inculcate
> 1st ed.   If you **inculcate** something such as an idea or an opinion in someone's mind, . . .
> 2nd ed.   If you **inculcate** an idea or opinion in someone's mind, . . .

or unnecessary repetitions:

bet
> 1st ed.   (def. 1) If you **bet** on a future event such as a horse race or **bet** someone an amount of money, you agree with someone an amount of money that they will give to you if the event happens in the way you have predicted, or that you will give to them if it doesn't. People sometimes bet on the result of a horse race or a sports match.
> 2nd ed.   (def. 1) If you **bet** on the result of a horse race, football match, or other event, you give someone a sum of money which they give you back with extra money if the result is what you predicted, or which they keep if it is not.

However, too simple definitions are not helpful, where necessary information, either linguistic or encyclopaedic, is omitted[1]:

hill
> 1st ed.   (def. 1) A **hill** is an area of land that is higher than the land that surrounds it, but not as high as a mountain.
> 2nd ed.   (def. 1) A **hill** is an area of land that is higher than the land that surrounds it.

---

1) Also compare: "A **mountain** is a very large raised part of the earth's surface with steep sides which are usually difficult to climb" (1st ed. def. 1); "A **mountain** is a very high area of land with steep sides" (2nd ed. def. 1).

4.1.2.6.   It is surprising that the 'Word Not In rule' (Landau 1984: 129) is sometimes broken even in *COBUILD*, which has been compiled almost totally by means of computers. For instance, to define *cocoa* the first edition used *cacao*, which was undefined. In the second edition, in which *cacao* is not included, the *cacao* has been replaced with "a tropical tree" in that entry.

However, at **HIV**, for instance, *immunodeficiency* is an undefined word. The lexicographer is unavoidably faced with the dilemma of whether or not to show a full spelling for abbreviations including such uncommon technical terms.[1]

4.1.3.   Defining Vocabulary The COBUILD series has not put any restrictions on the vocabulary used in explaining meanings (cf. Kojima et al. 1989: 96). "Carefully selected defining vocabulary", says the back cover of the second edition as one of the eight distinctive features. Has the new edition used restricted defining vocabulary?

The answer is "No". In fact, there is a clever explanation in the introduction[2]: " . . . a natural defining vocabulary with *most* words in our definitions being amongst the 2,500 commonest words of English" (italics added). Ms Fox admitted that they had not adopted a defining vocabulary system.[3] Theoretically, each lexicographer was allowed to use any word when necessary.

Is this 'cheating'? She asked me, however, how different *LDCE* is from *COBUILD*. Longman lexicographers are permitted to use virtually any word by simply printing it in small capitals, aren't they? She is right.

After all, the basic rule should be: Avoid using uncommon 'difficult' words when they are the key words in understanding the meaning of the word defined. Thus, *savoury* in the following excerpt would not cause serious trouble:

---

1) Like *COBUILD²*, *OALD⁵* has used the same undefined technical term, while *LDCE³* has avoided showing the full spelling of *HIV*. *CIDE* gives entry status to *immunodeficiency*.

2) *COBUILD²*, p. xviii.

3) Personal communication.

---

parsley
    COBUILD²: . . . used for flavouring or decorating savoury food.
    LDCE³: . . . used in cooking or as decoration on food

In this case *COBUILD* offers extra information by using it. Nevertheless, we have to say that 'difficult' words can easily make foreign learners feel uncomfortable.

4.2.   Separating and Ordering Different Senses

4.2.1.   'Superheadwords' The new edition has given up the 'one word, one entry' policy[1] of the original edition by introducing a new feature called 'superheadwords'.[2] It is welcomed because we were often "irritated"[3] by that strict policy.

Incidentally, the new edition is more 'user-friendly' with such entries as **new**, **parrot**, and **take**, where **news**, **parrot-fashion**, and **take** *a bath* are now all given headword status.[4]

4.2.2.   Separation of Senses and 'Sub-paragraphs' As mentioned in 4.1.1.1, senses of a polysemous lexeme, whether it is a single word or a set phrase, are now clearly divided from each other. This is a decision that will be welcomed by many users. When two or more meanings are felt not to be related to each other, whether etymologically related or not, they should be separated clearly in an EFL dictionary. A naïve user might otherwise wonder how similar "a lively folk dance" is to "a device that holds something in position",[5] for instance.

How we may know when a single lexeme has one meaning alone and when it has two or more quite discrete meanings is an annoying problem that linguists and lexicographers have been, and are, trying to answer. The first edition often allocated a separate paragraph or sub-paragraph to each

---

1) *COBUILD²*, p. x.

2) See 2.3.2–2.3.3 for a close examination.

3) *COBUILD²*, p. x.

4) The last is under a superheadword "used with nouns describing actions".

5) *COBUILD¹*, s.v. **jig**, def. 1.1 and def. 1.3, respectively. They are now under def. 1 and def. 3, respectively, in the second edition.

sense, even though each one is obviously related to other senses (e.g. **breathless**). It seems to be a basic rule of the second edition that related senses should be explained together in the same paragraph.

The lexicographer should take selectional restrictions into account when he explains related senses in a single paragraph. In the following entry this rule is carelessly broken[1]:

cockpit (2nd ed.)
> In an aeroplane or racing car, the **cockpit** is the part where the pilot sits.

Perhaps this is a rare example; one of the good features of the COBUILD dictionaries is its clear separation of senses according to selectional restrictions (and typical collocations) (cf. Kojima et al. 1989: 99–100). Here is such an example[2]:

die
> 1st ed.   **3** When things **die**, they function or burn more and more slowly and eventually stop completely.
> 2nd ed.   **4** When a machine or device **dies**, it stops completely, . . .
>   **5** When a fire or light **dies**, it stops burning . . .

4.2.3.   Order of Different Senses It was explained in the first edition how senses were arranged,[3] but the original edition was surprisingly inconsistent. According to the Introduction, colloquial expressions *I bet*, etc. were put third in the entry for **bet**, even though they occurred most frequently in the corpus. This was because *bet* in these phrases is not used as its "central, core meaning".[4] However, phrases like *of course* and *on the contrary* were listed first at the entries for **course** and **contrary**, respectively.

We looked for similar explanations in the new edition only to find very concise notes, which say, "**Phrases**: usually the last paragraph or paragraphs of an entry, before phrasal verbs . . . **Phrasal verbs**: in alphabetical

---

1)   This sense was separated into two sub-paragraphs in the first edition.
2)   Compare *COBUILD*² with *LDCE*³ and *OALD*⁵ in this respect.
3)   *COBUILD*¹, pp. viii–ix, p. xix. See also Kojima et al. (1989: 97).
4)   *COBUILD*¹, p. xix.

order at the end of an entry . . . "[1] The second edition seems to have obeyed the above principles fairly faithfully. Now *on the contrary* is listed fourth and fifth, and *of course* is given headword status.

However, there are exceptions. The phrases *regardless of* and *since time immemorial/from time immemorial* are still listed first under **regardless** and **immemorial**, respectively. This is perhaps because the second meaning of *immemorial* is usually used in a literary context (so it has an appropriate usage label) and the "of" in *regardless of* is regarded as a strong collocate.[2]

In fact, it is not at all unusual that 'central, core meanings' do not come first in the second edition (e.g. **bastard**, **bitch**, and **fuck**). There frequency is considered more important than semantic relationships between meanings, which we do not think an ideal decision from lexicographers of a foreign learner's dictionary.

<div align="right">(K. Nakamoto)</div>

5.   Usage
5.1.   Usage Labels

The COBUILD dictionaries indicate geographical and stylistic differences in terms not of 'labels' but of 'phrases' in accordance with the strict rule that meanings are all explained by full sentences. The new edition is no exception. In this review article, however, by 'usage labels' are meant these 'usage phrases'.

There is one respect in which the dictionary has been improved: now the dictionary provides the user with a list of geographical and style labels,[3] which contains much more information than the very brief guide in the first edition.[4]

According to this list, the second edition has employed two geographical labels[5] and 11 style labels.[6]

---

1)   *COBUILD*², p. xvi.
2)   Cf. *COBUILD*², p. x, p. xviii.
3)   *COBUILD*², pp. xx–xxi.
4)   *COBUILD*¹, p. xi.
5)   Namely, 'British English' and 'American English'.
6)   Namely, 'formal', 'informal', 'journalism', 'legal', 'literary', 'medical', 'offensive', 'old-fashioned', 'spoken', 'technical', and 'written'.

It has also used such labels as 'used showing approval/disapproval' to indicate the speaker's (or writer's) attitudes and feelings. These 'quasi-labels' are explained in another part of the Introduction.[1]

It is essential to list and define all the labels used in a dictionary. In fact, the first edition defined its usage labels, not in the introductory guides, but in the normal dictionary section as a part of the sense of the word used as a label. For instance, the label 'formal' was explained as follows (def. 1):

> **Formal** speech or behaviour is very correct and serious rather than relaxed friendly, and is used for example in official situations or when you are talking to someone important. In this dictionary, language of this kind is indicated by the use of the word 'formal' in definitions

The user might have found this explanation only by chance. Besides, 'formal', 'informal', and 'literary' were defined, but 'humorous', 'offensive', and 'old-fashioned' were not. The second edition is much more 'user-friendly' in this respect.

It seems that the new edition has dropped 'label modifiers' such as *very*, *fairly*, and *rather* that express varying degrees of formality, etc., thus making usage labels simpler[2]

	1st ed.	2nd ed.
eyeful	(def. 2) very informal	informal
accolade	fairly formal	formal
kin	rather old-fashioned	old-fashioned

We doubt the usefulness of these 'label modifiers' when they are used without any comparison. Thus simpler labels of the second edition are preferred.[3] However, the second edition has employed such modifiers as *mainly*[4] and *especially*.[5]

---

1) *COBUILD*[2], p. xxxvi.
2) Cf. Kojima et al. (1989: 104).
3) *CULD* is a pioneering work in which synonyms of different stylistic values are occasionally shown introduced by the phrases 'more/less formal than . . .', as in "*more formal than* **be sorry**" (s.v. **regret**) and "*less formal than* **telephone**" (s.v. **phone**).
4) E.g. "used mainly in American English" (s.v. **push-up**).
5) E.g. "used especially by children" (s.v. **wee** def. 2).

5.1.1. Stylistic Labels

5.1.1.1. As mentioned above, the current edition has used 'quasi-labels', the labels not listed in the front guide[1]:

piggy
> 1st ed.   (def. 2) . . . ; a rather childish use.
> 2nd ed.   (def. 1) . . . ; used by children

This kind of information about the 'users' of a particular word is often supplied as a part of a defining sentence:

pussy
> 1st ed.   (def. 1) A **pussy** or a **pussycat** is the same as a cat; used in informal English, often by children or when you are talking to children.
> 2nd ed.   (def. 1) Children or people talking to children often refer to a cat as a **pussy**.

These two explanations virtually contain the same amount of information (except that about formality; **pussycat** has a different entry in the second edition).

In the following entry, however, the latter explanation is considered better, even though "big robberies" is a very useful piece of information, because the user has to 'decode' the meaning of the "of" used in the former definition and because "often used of . . . " is NOT a usage label but a part of the meaning of the word defined[2]:

mastermind
> 1st ed.   (def. 2) The **mastermind** is the person who is responsible for planning and organizing a difficult or complicated activity; often used *of* criminals who plan big robberies. (italics added)
> 2nd ed.   The **mastermind** behind a difficult or complicated plan, often a criminal one, is the person who is responsible for planning and organizing it.

---

1) See *COBUILD*[2], pp. xx–xxi.
2) Here, too, 'use(d)' is very confusing. See 4.1.2.2.

5.1.1.2. The new label 'journalism' could have been replaced with similar explanatory notes mentioned in the preceding section, for instance, 'used (mainly) by journalists'.[1] In fact, all the field labels could be replaced with a simple phrase 'used (mainly) by . . . '.

If used alone the label 'technical' is not very useful for the foreign learner. It does not tell him in which field the headword is (mainly) used:

curvature[2]
> The **curvature** of something is its curved shape, especially when this shape is part of the circumference of a circle; a technical word.

In such cases, the COBUILD dictionary, both editions, usually specifies a particular field by means of the phrase 'in ...', as in "a technical term in music" (s.v. **fugue**).

However, there is a problem. Are *coda*, *full-back*, and *synergy* 'technical terms' in music, rugby or football, and business, respectively?[3]

> **coda** (def. 2) : In music, a **coda** is . . .
> **full-back**   : In rugby or football, a **full-back** is . . .
> **synergy**   : . . . ; used mainly by business people.

5.1.1.3. The helpful label 'spoken' could have been used much more generously. There are inconsistencies: **school kid, whoops** (EXCLAM), and **to cut a long story short** are all "used in spoken English", but **chap, whoopee** (EXCLAM), **hi** (CONVENTION), and **to be frank (with you)** are not. Shouldn't 'spoken' be given to any lexeme with the grammatical label 'EXCLAM' or 'CONVENTION'?

5.1.2. Geographical Labels The reviewers criticised the first edition for its poor coverage of American English and its bias towards British English.[4]

Now a huge corpus named 'The Bank of English' is accessible to the

---

1) See **wild child** in 4.1.2.3.
2) The same definition and label are given in both editions.
3) The excerpts are all from the second edition.
4) See Hausmann and Gorbahn (1989: 53), for example.

COBUILD lexicographers. Does it cover American usage fairly well?

Unfortunately, they admit that the corpus is still heavily dependent on British sources. The introduction says, "Although most of the sources are British, approximately 25% of our data comes from American English sources".[1] The percentage is obviously not high enough.

However, the COBUILD dictionary has been much improved in this respect. A random sampling test prompts the rough conclusion that the new edition marks more British expressions as such. In the first edition, American terms were often simply replaced by their unmarked British equivalents.[2] But now they are fully defined; typically, both British and American uses are defined in their entries, an American equivalent being given in the British entry, and vice versa[3]:

> **lift** (def. 7)
> In British English, a **lift** is a device that carries people or goods up and down inside tall buildings. The American word is **elevator**.
> **elevator**
> In American English, an **elevator** is device [*sic*] that carries people up and down inside buildings. The usual British word is **lift**.

This is an ideal method, except for its unhappy mistakes and inconsistencies, to indicate the two varieties of English. However, there remain lexicographic 'bugs' in the current edition, as in the original edition. For instance, if you look up **trolley** (def. 1), you know it is a British word and that its American equivalent is *cart*. However, if you happen to look up **cart** (def. 4), you will never know its British equivalent is *trolley*.[4] Another example is **subway** (def. 1). An opposite case is **the box** (s.v. **box** def. 6), where its American equivalent, *the tube* (cf. s.v. **tube** def. 5), is not shown, while the latter entry does give that British expression. The entry for **bank holiday** tells the user that its American term is *national holiday*, but it is

---

1) *COBUILD*², p. xii.
2) See **hood, muffler**, and **streetcar**, for example. Cf. Kojima et al. (1989: 105).
3) It is obvious, however, that the following two difinitions were writting by different lexicographers. Don't elevators carry goods?
4) It is of course good that both *trolley* and *cart* are defined in each entry.

not included in the dictionary.[1] It is hoped that both British and American entries will be checked thoroughly by the COBUILD lexicographers.

About 5% of the data stored in 'The Bank of English' comes from sources of "other varieties of English — such as Australian and Singapore".[2] No such geographical labels as 'Australian' and 'Singapore' are used, however, perhaps because the data is not large enough. This is all right at least for Japanese learners of English, who definitely need much more information about American English.

5.2. Usage Notes

Unlike other EFL dictionaries (e.g. *LDCE* and *OALD*), neither the first nor the second edition of *COBUILD* contains 'Usage Notes' as such. Information about usage is usually given under usage labels (mentioned above) and/or a 'PRAGMATICS' label. Warnings are sometimes given before or after the definition, particularly in entries for four-letter words.

By avoiding usage columns, however, the COBUILD dictionary suffers a serious disadvantage: it cannot explain semantic, grammatical, and collocational differences between synonymous lexemes. Synonyms are scattered all through the alphabetical word list. Occasionally, a synonymous word is given in the Extra Column being introduced by the equal symbol (=). However, it is not very helpful. The dictionary does not tell the user about the differences between *travel*, *journey*, *trip*, and *voyage*, for instance.

(K. Nakamoto)

6. Examples

The Introduction to the new edition says that the examples first and foremost "show the characteristic phrasing round the word" — in other words, they give "a full display of the usage of the word".[3] This is possible because the COBUILD style of defining makes definitions clear enough in themselves.

---

1) Another example is **double cream** — *heavy cream.*
2) *COBUILD*[2], p. xii.
3) *COBUILD*[2], p. ix.

Several questions should be posed: Are examples really 'authentic' so that they reflect typical uses of a headword? Which is more important in a foreign learner's dictionary, 'authenticity' or 'understandability'? Are definitions so clear that the user does not have to rely on examples as a short cut to the meaning? How indeed are sentence definitions different from examples?

6.1. Authenticity

The first edition's examples were criticised, often in a very negative tone, by its reviewers, among whom were Hausmann and Gorbahn (1989: 46) who listed seven "weaknesses" in them. Their criticism was directed particularly at the *COBUILD*'s claims that the dictionary represents authentic or 'real' English. According to them, authentic examples "often sound strange", "detract from the actual definition of a word", and "reflect a very idiosyncratic use of the English language" (Hausmann and Gorbahn 1989: 46).

"All of the examples in this book [i.e. the second edition] are newly selected from The Bank of English".[1] Is this because the examples in the previous edition were criticised, or because they were totally unsuitable or even incorrect?

Of course, not. The dictionary explains that the examples were all replaced simply because "this is a completely new edition".[2] Ms Fox said that by replacing older examples with new ones the new edition can present different data and thus becomes a supplementary reference book.[3] What a good idea! What a wise commercial strategy!

The 'authentic' examples shown in the second edition must be more reliable; the corpus on which it depends contains over 200 million words. The corpus must be large enough to show "typical grammatical patterns, typical vocabulary, and typical contexts".[4]

It should be noted that occasionally "very minor changes" have been

---

1) *COBUILD*[2], p. ix.
2) *COBUILD*[2], p. xxii.
3) Personal communication.
4) *COBUILD*[2], p. xxii.

made to the examples directly taken from the corpus.[1] However, the present reviewer's rough impression is that the COBUILD examples are still rather 'difficult'.

6.1.1. 'Difficult' Words Used in the Examples

As in the defining sentences,[2] if any 'difficult' word, whose meaning is 'difficult' for the learner to understand, is used in an example, it will become a 'difficult' example. This is particularly true for cases where the word is a 'key word', the meaning of which needs to be known already for an understanding of the whole meaning of the example. No example could be used by the learner for encoding purposes unless it is correctly understood. For instance, *lucrative* is a key word for understanding the following example: *It was quite a lucrative sideline* (s.v. **sideline**, def. 1).

It is unrealistic to restrict the vocabulary used in the examples (as *LDCE*[1] did), but it is equally unrealistic NOT to restrict it in a foreign learner's dictionary. We believe that editorial adaptation, "very minor" or not, is unavoidable even if this spoils the authenticity of the examples. The lexicographer would otherwise have to look for the best possible example often from almost untolerably long concordance lines.

Uncommon proper nouns should be avoided, unless they are defined (or explained) in the dictionary:

Scholars have debated whether or not Yagenta became a convert (s.v. **debate**, def. 3)

The flat-pack units are by Gower kitchens (s.v. **flat pack**)

In the above examples, *Yagenta* and *Gower* can easily distract an innocent learner's attention.

6.1.2. Unclear Contexts

The lexicographer will have to run the risk of leaving unclear the context of a particular example by taking a short passage directly from the corpus. This was an aspect many reviewers criticised after the publication of the first edition.[3] The third example for

---

1) *COBUILD*[1], p. xv; *COBUILD*[2], p. xxii.

2) See 4.1.3.

3) See Hausmann and Gorbahn (1989), for example.

---

motive, which is cited in the Introduction[1] with its concordance, goes: *The doctor's motive was to bring an end to his patient's suffering.* The doctor's motive for what? It is the user who has to make up for the missing information.[2]

Pronouns, especially *this, that*, and their plural forms, should be avoided, unless their antecedent is specified. What does "this" refer to in the following example: *Me and Ben should sort this out man to man* (s.v. **man**, def. 20).

6.2. Examples and Defining Sentences

The problem discussed at 4.1.2.3 can be looked at from the opposite angle: How different are the *COBUILD*'s definitions from examples?

Here take **amnesia** for example[3]:

> *COBUILD*[2]: If someone is suffering from **amnesia,** they have lost their memory. *People suffering from amnesia don't forget their general knowledge of objects.*
>
> *OALD*[5]: partial or total loss of memory: *suffer an attack of amnesia.*

The noun *amnesia* means "(partial or total) loss of memory", as *OALD*[5] shows. On the other hand, *COBUILD* gives the meaning of the PHRASE *to be suffering from amnesia*, instead of the meaning of the WORD *amnesia*. The first half of the definition could be regarded as an example that shows "characteristic phrasing round the word". Indeed the same collocation is shown in the example sentence, which, in turn, could be regarded as a part of the definition.[4]

Here is a much worse example. *COBUILD* does not tell its user what the noun **summary** means; the first paragraph explains the meaning of the phrase **in summary**, where we find a wonderfully circular definition,

---

1) *COBUILD*[2], p. xxiii.

2) Perhaps the doctor practised euthanasia on the patient who had been suffering from a terminal illness.

3) Incidentally, the new edition has abolished the example marker 'EG' that separated examples from the definition in the former edition.

4) Compare: If someone is suffering from **amnesia**, they have lost their memory, but they don't forget their general knowledge of objects.

and the second explains the adjectival use, which has a different meaning from the nominal use. To make matters worse, *summary* is also used to define its verbal form **summarize**: If you **summarize** something, you give a summary of it. How can the learner understand the examples presented there if he does not know the meaning of *summary*?

Does the COBUILD style of defining always explain meanings clearly enough? Isn't too much emphasis put on the encoding function of the defining sentence sacrificing its essential decoding function?

6.3. Other Features and Problems

6.3.1. Grammatical Patterns Unlike the first edition where grammatical labels were given to the headword itself, the second edition gives the examples "in the same order as the patterns shown in the Extra Column".[1] This is virtually the same as showing grammatical labels before each example (as in *OALD*[5]). The new helpful system is mainly applied to verbal entries. In entries for nouns, adjectives, and adverbs, however, grammatical labels are basically given to the headword itself, and grammatical patterns shown in the Extra Column are not necessarily exemplified. For instance, to the second sense of **immersion** is given a grammatical label 'N-UNCOUNT: oft N *in* n'. The example sentence supplied shows the uncountable usage, but it does not exemplify the use with the preposition *in*.

6.3.2. Entries or Paragraphs with No Examples "Nearly all the words and meanings in the dictionary have at least one example. The main exceptions are concrete nouns . . . , and a few other words where an example would add nothing to the information given in the definition", says the Introduction.[2] It is a shame that restricted space often prevents the lexicographer from offering examples, even when he has a large reliable corpus from which typical and useful collocations can be taken.

Carefully chosen examples can do more than lexicographers would ex-

---

1) *COBUILD*[2], p. xxii.

2) Ibid.

pect, even at entries for very popular concrete nouns. The second edition does give three examples to the first sense of **dog**: *Outside, a dog was barking*, *The dog growled again*, and *The British are renowned as a nation of dog lovers*. However, no example is shown with **cat** (def. 1), perhaps because the British like dogs better than cats! The lexicographer could have given it such examples as *Cats miaow when they are unhappy, purr when they are happy*, *We'll have to get someone to feed the cat when we're away on holiday*, and a *pet/stray cat*, the first example being taken from *COBUILD*[2] (s.v. **miaow**) and the others from *CIDE* (s.v. **cat**). We believe these examples could add something to the information given in the definition.

In *COBUILD*, however, the defining sentence often gives useful information. The problem, after all, is in whether the user can 'decode' that information. For instance, from the latter half of the definition of **cat** (" . . . Cats are often kept as pets") a useful collocation *keep a cat as a pet* could be found.

On the other hand, every lexicographer should consider what examples are really useful for the potential user. How useful is the following example?: *I want an apple* (s.v. **apple**, def. 1).

There are obviously more than "a few" words (other than very common nouns) to which no example is given. Take, for example, terms relating to diseases again. Since no example is provided, the user can learn which verbs typically collocate with nouns like **AIDS, Alzheimer's Disease, diabetes, haemophilia, mumps, tuberculosis,** and **typhus** from neither examples nor definitions (cf. 4.1.2.3).[1]

<div align="right">(K. Nakamoto)</div>

7. Grammar

7.1. The Types of Grammar Information

Grammar information in *COBUILD*[2] are of three types: the word class of the word, restrictions or extensions to its behavior and the patterns in

---

which the word most frequently occurs (p.xxiv). The word class appears in capital letters and the rest of the information mainly appears in small letters. The words in italics are words that occur in the structural pattern. The system of providing the information in the extra column is also employed in the new edition.

Synonyms, antonyms and superordinates also appeared in the extra column in *COBUILD*[1], but superordinates are not referred to in *COBUILD*[2].

In *COBUILD*[1], users had to refer to the special entries which were arranged alphabetically among other word and phrase entries in order to decipher the grammatical notations properly. However, in the new edition, all grammatical codes are explained in the front matter (p. xxiv–xxxiii) of the dictionary, as in many traditional dictionaries. This system is more user-friendly compared with that in the previous edition, for the users can consult the grammar information without looking for a certain grammatical term which appears on various pages of the dictionary.

According to the editors, efforts were made in order to simplify and improve the grammar notes in *COBUILD*[2] (p.xi). Some of their efforts will be examined in the following sections.

7.2. Structural Patterns

As for structural patterns, the patterns of the new edition are more minute and precise than those of *COBUILD*[1]. Take the verb *declare* for example. The grammar notes for *declare* 4 in *COBUILD*[1] show that the verb is frequently used in patterns of three types: $V + O$, $V + O + C$, $V + REPORT\text{-}CL$. In *COBUILD*[2], the same patterns are indicated more precisely as $V\ n$, $V\ n\ adj$, $V\ that$ respectively. In this edition, the last item in the second pattern, which functions as a complement, is explicitly shown as an adjective. As for the third pattern, the new code *that* is used instead of the code *REPORT-CL* to make it clear that the clause which may follow the verb is a *that* clause. In fact, the code *REPORT-CL* is no longer used in *COBUILD*[2] for the code was polysemous in that it represented clauses beginning with *that*, *wh* words, *if* and *whether*.

It was pointed out in the review of *COBUILD*[1] (Kojima *et al*, 1989:128)

that sometimes the patterns in the extra column did not correspond to the examples given after the definitions of words or phrases. In *COBUILD*[2], however, the patterns are given next to the corresponding examples, which may be regarded as an improvement in this edition. In the review, the verb *bet* was given as an example to illustrate the incongruity between the patterns in the extra column and the example sentences given after the definition. In *COBUILD*[1], the second example shows that there is a possibility that a *that* clause follows the verb, but such a pattern does not appear in the extra column. The definition also suggests the intransitive use of the verb, but an example which illustrates the use is not given after the definition. In *COBUILD*[2], however, the intransitive use is illustrated by the first example, and the pattern in which a *that* clause follows the verb also appears in the extra column.

7.3. New Grammar Notations

7.3.1. In order to describe the possible structural patterns precisely, the number of grammar notations has been increased to 129 in the new edition. Not only the number of word classes but also that of the words and abbreviations used for the description of patterns has been increased. The meanings of notations used in patterns seem rather easy to guess at, except *brd-neg* (broad negative), which may be unfamiliar to the users. As for word classes, however, there are some new codes the meanings of which are rather difficult to understand without referring to the explanation in the front matter.

7.3.2. *COBUILD*[1] was the first dictionary to use the category of ergative verbs in classifying verbs. Later it was recognized that there were phrasal verbs which could be regarded as ergative verbs (*CCEG* p.167). Those phrasal verbs are indicated by the code *ERG* in *COBUILD*[2]. Take the phrasal verb *block up* as an example. In *COBUILD*[1], the notation for the phrasal verb is just *PHRASAL VB*. In *COBUILD*[2], however, the code *ERG* is attached to the notation *PHRASAL VB*. The example sentences given after the definition prove the ergativity of *block up* (p.167):

(1) *Powdering a sweaty nose will only block up the pores and make the skin uncomfortable.*

(2) *With this disease the veins in the liver can block up, and all sorts of damage follows.*

7.3.3. The code *V-LINK*, which stands for a link verb, is another newly introduced code in *COBUILD*². A link verb is a verb which links a subject and a complement (*CCEG* p.173). The typical examples such as *be* and *become* are indicated only as *V* in *COBUILD*¹ while they are indicated as *V-LINK* in *COBUILD*². It is noteworthy that some phrasal verbs are regarded as link verbs in the new edition. The phrasal verb *turn out* is an example (p. 1801):

> *It's turned out nice again.*

7.3.4. In *COBUILD*², a type of verb which has both the features of an ergative verb and a link verb is recognized. They are called ergative link verbs and the code *V-LINK-ERG* is used to represent them. The verb *turn* is an example. It behaves like a link verb when it connects a subject and a complement in a pattern such as *V colour: The sea would turn pale pink and the sky blood red* (p.1799). The verb, however, can also occur in a *V n* pattern in which an ergative verb can appear: *Her contact lenses turned her eyes green* (ibid.).

7.3.5. *COBUILD*¹ was the first dictionary to recognize reciprocal verbs as a category in classifying verbs. In *COBUILD*², the existence of verbs which behave both like ergative verbs and reciprocal verbs is recognized. They are called ergative reciprocal verbs and indicated by the code *V-RECIP-ERG*. The verb *combine* is an example. The verb has the following patterns: *X and Y combined, X combined with Y,* and *Z combined X and Y.* It behaves like a reciprocal verb in that it can occur in the second pattern in which the two factors are involved in the process. At the same time, it can be regarded as an ergative verb, for it appears in the first and the third of the patterns (Francis and Sinclair, 1994:199).

As with ordinary verbs, some phrasal verbs are ergative reciprocal verbs. Take *pair up* as an example. There are three examples after the definition (p.1192):

(1) *They asked us to pair up with the person next to us and form teams.*
(2) *Men and teenage girls pair up to dance.*

(3) *Smokers and nonsmokers are paired up as roommates.*

The example (1) is a reciprocal pattern, while the example (2) shows the intransitive pattern. The third example can be rewritten as follows:

(4) *They paired up smokers and nonsmokers as roommates.*

The second and the fourth examples show the ergative relationship. Therefore this phrasal verb is an ergative reciprocal verb. The example (4) can again be rewritten as (5):

(5) *They paired smokers up with nonsmokers as roommates.*

Therefore the typical patterns in which an ergative reciprocal verb occur are *pl-n v, v with n, v pl-n, and v n with n.*

7.3.6. As with ordinary verbs and phrasal verbs, some phrases are recognized as ergative phrases in *COBUILD*². The pair *to turn one's stomach/ one's stomach turn* provides an example. Compare the following (p.1642):

(1) *The true facts will turn your stomach.*
(2) *I saw the shots of what happened on television and my stomach just turned over.*

Since the phrase can occur in both *V n* and *V* patterns, it is classified as *PHR-ERG* in *COBUILD*².

7.3.7. The concept of ergativity is not the only concept which has been extended to phrases in *COBUILD*². *COBUILD*² considers some phrases as reciprocal phrases when they contain a verb, and behave like a reciprocal verb. The patterns in which *PHR-RECIP* occurs are *pl-n PHR* and *PHR with n.* The phrases *settle accounts/ settle one's accounts* are examples (p. 13):

(1) *Their sleep is regularly disturbed by the sound of gunfire as criminal gangs settle their nightly accounts.*
(2) *Germans could finally settle accounts with the British.*

7.3.8. A phrasal modal is another new categorty in *COBUILD*². It is a phrase which occurs before the infinitive form of a verb and behaves as a modal. It is indicated by the code *PHR-MODAL*. The phrases *used to* and *would rather* serve as examples. In *COBUILD*¹ the phrase *used to* is indicated as a semi-modal but *would rather* is not even recognized as a phrase.

7.3.9. A phrase which behaves like a preposition is classified as a phrasal preposition in *COBUILD²*. The phrases such as *in spite of* and *according to* serve as examples. Both phrases are indicated only as *PREP* in *COBUILD¹*.

7.3.10. A phrasal coordinating conjunction is a conjunction which is combined with another word or a group of words and used in the form of a phrase. It is indicated by the code *PHR-CONJ-COORD*. Similarly, a subordinating conjunction which is combined with another word or a group of words and used in a form of a phrase is called a phrasal subordinating conjunction. The code *PHR-CONJ-SUBORD* is used to indicate it. In *COBUILD¹*, the members of the two categories are classified only as coordinating conjunction and subordinating conjunction respectively. An example of the former is *or else* and an example of the latter is *just because*:

(1) *Evidently no lessons have been learnt or else the government would not have handled the problem so sloppily.* (p.537)
(2) *Just because it has a good tune does not mean it is great music.* (p.135)

7.3.11. The code *PHR-* is one of the key codes among the grammar notations used in *COBUILD²*. The code indicates efforts have been made to highlight the fact that phrases play an important part in English grammar. The attachment of the code *PHR-* to a certain word class sometimes provides important information for the user. Take the phrase *be able to* as an example. In *COBUILD¹*, the notation for this phrase is *ADJ CLASSIF* since it is classified according to the word class of the adjective *able*. It is not recognized that the word *able* behaves like a modal when combined with the verb *be* and used before a *to*-infinitive. In *COBUILD²*, however, the fact that the adjective is used in the form of a phrase and behaves like a modal is explicitly indicated by attaching the code *PHR-MODAL* to it.

7.4. The Word Classes

The number of word classes in *COBUILD²* is very large compared with other EFL dictionaries. In fact *COBUILD²* classifies the entries into 73 categories while there are only 17 word classes in *LDCE³*. This is because

COBUILD² considers the restrictions or extensions to the words' behavior as a part of word class. Take the noun *committee* as an example. This noun, as we all know, can take either a singular verb or a plural verb after it. In *LDCE³* this fact is indicated in square brackets since it is considered as an extension to the word's behavior. In *COBUILD²*, however, the noun which can take either a singular verb or a plural verb is classified into an independent category. Therefore the noun *committee* is classified as a collective count noun. Another example is the noun *injustice*. This noun can be used in three forms: *injustice, an injustice*, and *injustices*. This fact is indicated as [C,U] in *LDCE³* while *COBUILD²* classifies the noun as a variable noun. A variable noun is defined as a noun which combines the behavior of both count and uncountable nouns (p. xxvii).

However, this system of incorporating the restrictions or extensions to words' behavior into word classes does not seem to be user-friedly, because there is no guarantee that the users read the explanation of grammar codes in the front matter before using the dictionary. The users, therefore, may not always be able to draw out the information which concerns the behavior of the words from the word class codes. In the case of the noun *injustice*, the user is expected to find out that the singular form of the noun can be used either with or without the determiner by looking at the word class code *N-VAR*. The user, however, may not be able to figure out that the code stands for a variable noun, and even if they did, they may not be able to find out the fact that the noun has two singular forms since the word *variable* is not self-explanatory. Sometimes the information concerning the behavior of words can be inferred from the examples. However, the example sentences will be of no help in this case, for the singular form *an injustice* is not illustrated in the example sentences. Is it taken for granted that the users memorize all the meaning of word class codes in order to use *COBUILD²* effectively?

As we look at the list of word class codes in *COBUILD²*, we cannot help asking the question: are they all necessary? The number of the codes just overwhelms us. Take the nouns as an example. The nouns are divided into 16 categories. Among them, there are five types which has the code *-COLL* attached. As mentioned in the above section, it may be difficult for

the users to infer from this code that this type of noun can take either a singular or plural verb. It may, therefore, be more user-friendly to provide the information in explicit notation instead. Then can the use of the code -COLL still be justified? After all what is the advantage of encoding the grammatical features in word class codes?

7.5. The Disadvantage of Codes and Abbreviations

The reason why COBUILD[2] uses a large number of codes and abbreviations is that the dictionary sets out the grammatical information in the extra column separate from the definitions. The codes and abbreviations are useful since they do not take up much space; a large amount of information can be set out economically in a limited space. However, the disadvantage is that not all codes and abbreviations are self-explanatory. The users give up trying to decipher the codes and abbreviations when they are not understandable at the first sight, and eventually ignore the information conveyed by them. This is a cruel fact that every lexicographer has to keep in mind.

7.6. The Possibility of Restructuring the Grammatical Information

The possibility should be considered whether some grammatical information — especially the information about the restrictions and extensions to the behavior of words — may be incorporated in the definitions just like the information on style and usage. Then the information can be indicated more explicitly without using codes and abbreviations that are difficult for the users to decipher. The possibility of the users' ignoring the grammatical information should not be overlooked, since there are users who use the dictionary to avoid making grammatical mistakes in their linguistic performance.

It is expected that the grammatical notation of COBUILD[2] will be improved and simplified even more in the next edition in pursuit of user-friendliness.

(R. Takahashi)

8. Pragmatics

8.1. Introduction

8.1.1. COBUILD[2] gives a special label ⟨PRAGMATICS⟩ in the Extra Column to notify readers that a given word is provided with some pragmatic information. Reasons for the inclusion of pragmatics are explained in the Introduction and in the Guide to the Dictionary Entries:

> Many uses of words need more than a statement of meaning to be properly explained. . . . This aspect of language is very important, and easy to miss. This is where the language is giving added meaning.[1]

> Different languages use different pragmatic strategies. In order to use a language effectively, and be successful in achieving your goals, you need to know what the pragmatic conventions are for that particular language. It is therefore important that learners of English are given as much information as possible about the ways in which English speakers use their language to communicate.[2]

8.1.2. Pragmatic information in the dictionary, we believe, is very useful for advanced readers and writers of English in our country as well as for beginners. Quite often, Japanese learners do not know whether 'easy-going' in 'He is easy-going.' approves or disapproves him. In Japan, where hard working is thought much of, an easy-going attitude sounds negative rather than positive. The Japanized word for 'easy-going' implies laziness and/or irresponsibility. Pragmatic information of this sort is valuable because 'different languages use different pragmatic strategies' and therefore in many cases, learners cannot draw an analogy from their own native language.

8.1.3. Pragmatic information is not a new feature to the dictionary. According to the Preface, 'COBUILD has always had a lot of information on pragmatics in its pages, but we have not previously drawn attention to it except in the case of insults, swear words, and things like that.'[3]

Other learner's dictionaries also have many pieces of pragmatic information

---

1) Introduction xi.
2) Pragmatics xxxiv.
3) Preface xi.

tion. *LDCE²*, for example, gives pragmatic information in the definition as well as in Usage Notes and in Language Notes.

8.2.   Comparison of Older Editions

8.2.1.   In this section, we will compare *COBUILD¹*, *LDCE²*, and *OALD⁴* in order to survey how pragmatic information is given in these dictionaries. Then we will compare *COBUILD¹* with *COBUILD²*, together with other revised learner's dictionaries to examine what improvements have been made.

A short comparison of three dictionaries will clarify the differences in treatment of pragmatic information and how they describe it. (Definitions are either omitted or simplified.)

absolutely

LDCE²   Usage 1 **Absolutely** is often used to give more strength to following adjectives or verbs which are already very strong.

OALD⁴   4 (used to give emphasis)

COBUILD¹   **1 Absolutely** is used **1.1** to emphasize what you are saying, especially when you are expressing an opinion.

of course

LDCE²   Usage **Of Course** (= certainly) is a polite way of agreeing and showing willingness to help, in reply to requests such as. . . . But **of course** is not polite in reply to a question asking for information:

OALD⁴   (no pragmatic information)

COBUILD¹   **1** You say **of course** when you are briefly mentioning something that you expect other people already realize or understand, or when you want to indicate that you think they should realize or understand it.

3 You say **of course** as a polite way of saying yes, of giving permission, or of agreeing with someone.

if (if I were you)

LDCE²   5 (used when giving advice)

OALD⁴   10 **if I were you; if I was/were in your shoes/place** (used to introduce a piece of advice to sb)

COBUILD¹   14 You say '**if I were you**' to someone when you are giving them advice.

8.2.2.   Of the three dictionaries, *LDCE²* gives most information on pragmatics. In the essay on Pragmatics and the Dictionary,[1] Professor Geoffrey Leech and Dr Jenny Thomas explain the importance of pragmatic information in a dictionary for learners of English.

Three means that *LDCE²* adopts to capture guidelines of pragmatic usage is: 1 Usage Notes, 2 Language Notes, and 3 Comments and examples within the entries for individual words.[2] Language Notes on Addressing People, Apologies, Criticism and Praise, Invitations and Offers, Politeness, Requests, Tentativeness, and Thanks concentrate mainly on pragmatics.

LDCE² gives labels such as *apprec, derog, euph, humor,* and *pomp* to show that the word suggests a particular attitude. Parentheses are also used before the definition to give another kind of pragmatic information.

8.2.3.   Pragmatic information in *OALD⁴* is less conspicuous. Like *LDCE²*, it gives labels such as *approv(ing), derog(atory), joc(ular), euph(emistic),* etc. in parentheses. Notes on Usage sometimes include pragmatic information,[3] but unlike *LDCE²*, *OALD⁴* pays no special attention to pragmatic information in special columns or pages.

8.2.4.   In *COBUILD¹*, information about language use is included within the entry for each word. There is no special column or parentheses that explains information on language use. All pragmatic information is inseparably incorporated in the definition of individual words. One major field of information on language use concerns the attitude of the speaker towards the words: **avaricious, bookish, kid's stuff, prissy,** etc. are words '(often) used showing disapproval.' Words like **ideally 2, principle 1** are 'used showing approval.' There are also Notes like 'offensive word, rude word, swear word,' etc. in the dictionary. Words and phrases used for apologizing, asking, inviting, thanking, warning, etc. are so described just

1) See F12.

2) See F13.

3) Refer to the Note On Usage of shall, may, etc.

the same way the meanings of the word are given: **wish 6** People sometimes say '**I don't wish to** be rude,' '**I don't wish to** interrupt,' etc. as a way of apologizing or of warning you before they say something which they think might upset, worry, or annoy you. There are words and phrases used for functions such as discourse organizers, emphasis, etc.: **way 25** You say **by the way 25.1** when you add something to what you are saying, especially a question or piece of information that you have just thought of.

8.3.   Comparison of New Editions

8.3.1.   Contrary to the policy adopted by *COBUILD²*, *LDCE³* has reduced a considerable part of its information on pragmatics. It has totally abandoned the Language notes. Usage notes on language use are much reduced. Instead, some are incorporated in the entry. Pragmatic information in the entry formerly given in parentheses remains (but without parentheses). *LDCE³* gives no reason why it has abolished Language Notes and has reduced other pragmatic information. It may have been deleted to create space for new information such as frequency, new phrases and collocations. Or, there might be some more radical reasons. The editors and lexicographers might have found that there is little significance in incorporating such pragmatic information in the dictionary. Anyway, it is a loss.
8.3.2.   *OALD⁵* has not changed its policy on pragmatics. Some information in parentheses is enriched but much remains essentially the same. One page of the language Study is devoted to Polite expressions.

8.4.   Pragmatic Information in the New Edition of *COBUILD*

8.4.1.   The New Edition of *COBUILD* has greatly enriched pragmatic information. The way *COBUILD²* has employed it is:

(a) to show this (pragmatic information) in the extra column with the word PRAGMATICS and
(b) to include additional information in the definition about how, when, and why the word or expression is used.

Our survey sampling on 196 pages (10% of the dictionary pages) shows that there are 2.2 pragmatic labels on every page. This means there are about 4300 pragmatic labels in the dictionary. *COBUILD²* classifies prag-

matic information into six major types. They are Functions, Discourse organizers, Speaker/hearer relationship, Attitudes and feelings, Emphasis, and Expressing certainty and uncertainty.[1] Although some information such as attitudes and feelings are incorporated in the definition, the dictionary does not provide any overt distinction between most of these pragmatic information types. So readers must decide by themselves what sort of pragmatic information is provided in each case.
8.4.2.   Dictionary definitions that give various types of pragmatic information typically begin the defining sentences as follows.

Functions:
> You can say '**I wonder**' if you want to be very polite when you are asking someone to do something, or . . . / If you tell someone to **watch out**, you are warning them to be careful, because . . .

Discourse organizers:
> You use **added to this** or **added to that** to introduce a fact that . . . / When you are talking, you can say that something **brings** you to a particular point in order to indicate that you have now reached that point and are going to talk about a new subject.

Speaker/hearer relationship:
> People sometimes say **sir** as a very formal and polite way of addressing a man whose name they do not know or a man of superior rank. / You call someone **darling** if you love them or like them very much.

Attitudes and feelings:
> If you refer to a group of people as a **charmed circle**, you disapprove of the fact that . . . / If you describe something as **masterly**, you admire it . . .

Emphasis:
> You use **need** in expressions such as **I need hardly say** and **I needn't add** to emphasize to the person you are talking to that . . . / If you describe something as **perfectly** good or acceptable, you are emphasizing that . . .

Expressing certainty and uncertainty:

---

1)   See p. xxxiv–xxxvii for details.

You say 'Not that I know of' when someone has asked you whether or not something is true and you think the answer is 'no' but you can not be sure . . . / You can use far in expressions like 'as far as I know' and 'so far as I know' and 'so far as I remember' to indicate that you are not absolutely sure of . . .

8.4.3.   A large amount of information on functions, discourse organizers, speaker/hearer relationship, and certainty and uncertainty originates in *COBUILD*[1] (information on the emphatic use may be new). **Speaking**, for example, has three pragmatics labels in *COBUILD*[2]: 1 **speaking as** a parent/a teacher . . . to indicate that the opinion you are giving is. . . , 2 **speaking of** . . . as a way of introducing a new topic which . . . 3 **generally speaking/technically speaking** to indicate the range or relevance. . . . Similar information is given in the definition of *COBUILD*[1]: 1 You use **speaking** with an adverb when you are defining the way you are describing something, **2.1** to indicate what your position or viewpoint is in what you are saying, **2.2** to introduce a new topic, by . . .

Many new pieces of pragmatic information are added in *COBUILD*[2]. Just a few examples (in a simplified form) will suffice for the illustration: be of **advanced** years (you are saying in a polite way), **advisable** (you are suggesting), **blackmail** (you disapprove), someone **jumps to a conclusion** (you are critical), **in broad daylight** (you are expressing your surprise), etc.

8.5.   For Improvement
8.5.1.   The presentation of pragmatic information in the way above stated may be not very helpful for the learners because the distinction between the meaning and the use of the word is unclear. For instance, the twentieth sense of **with** is defined as 'If someone says that they are **with** you, they mean that they understand what you are saying.' This sense is marked with a pragmatic label, but the pragmatic information implied here is not easy for learners to figure out. They may be not able to tell which part of the explanation is the lexical meaning of the word and which part is the pragmatic information. If the expression implies 'encouraging,' the dictionary should note that in the text.

The distinction between the lexical meaning of the word and the pragmatic information is very important for the dictionary, because the lexical meaning is inherent in the word while in many cases, the utterance meaning results from the context. Thus, the learners will wonder if the uncertainty expressed in the second sense of **think** is a part of the lexical meaning or derived from the use of the word: If you say that you **think** that something is true or will happen, you mean that you have the impression that it is true or will happen, although you are not certain of the facts.

To distinguish pragmatic information from the definition, conventional parentheses may be an easy way of doing so. Another easy way is to give information after the definition, separated by the semicolon as in 'Someone who is **balanced** remains calm and thinks clearly, even in a difficult situation; used showing approval.' This is the method both *COBUILD*[1] and *COBUILD*[2] have adopted for words that have a functional use such as in conveying the speakers' viewpoint. We believe all pragmatic information should be explicitly explained, i.e. what pragmatic information each illustrative sentence conveys.

8.5.2.   Some pieces of information on pragmatics are inconsistent. Words used for addressing like **darling, madam, sonny, sir,** etc. have pragmatic labels, but there are similar words used for addressing but have no label: **daddy, mummy, honey,** etc. **Push in** has a pragmatic label that says that it is used showing disapproval, but **push around** has no label. The word is explained as 'If someone **pushes** you **around**, they give you orders in a rude and insulting way' with an illustrative example: *We don't like somebody coming in with lots of money and trying to push people around.* The third definition of **rotten** is 'If you describe someone as **rotten**, you are insulting them or criticizing them because you think that they are very unpleasant or unkind.' Don't these words deserve pragmatic labels?

8.5.3.   Before concluding this section, we have to consider what kind of pragmatic information should be incorporated in the dictionary, and what should be left out. If pragmatics derives from each sentence rather than the word/phrase in the sentence, then should a dictionary give such information? The eleventh definition of **easy**, for example, has a pragmatic label: If you tell someone to **go easy on**, or **be easy on**, a particular person, you

are telling them not to punish or treat that person very severely. *'Go easy on him,' Sam repeated, opening the door.* . . . *Be a little easier on yourself and enjoy yourself more.* . . . From the definition and the illustrative examples, we can assume that pragmatic information implied here is 'advising.' Obviously, this is not the case in the next illustration given in the same paragraph. *This agency has been far too easy on the timber industry over the years.* This second sentence may have some other pragmatic information such as 'criticizing.' Is it the phrase **go easy on/be easy on** or the sentence itself that has the extra meaning? If the sentences that have pragmatic information should be noted in the dictionary, where can we stop? Professor Geoffrey Leech and Dr Jenny Thomas write that 'the very same words (Is that your car?) can be used to complain, to express admiration, or to express disapproval.'[1]

Generally, sentence forms such as imperatives and quasi-imperatives have special import. *Longman English Grammar*, for example, lists nine common uses of the imperatives. They are: 1 Direct commands, requests, suggestions. 2 Warnings. 3 Directions. 4 Instructions. 5 Prohibitions. 6 Advice. 7 Invitations. 8 Offers. 9 Expressing rudeness.[2] Imperatives including the above stated *go easy on, be easy on*, and *take it easy*, etc. addressed for the benefit of the person spoken to (or for the third party) usually imply 'advice, offers, suggestions, and invitations,' the inference deduced from the so called Politeness Principle.[3]

Pragmatic information of this sort may be better explained with its context on special pages as in Language Notes in *LDCE*[2], so that readers can grasp the speakers' intention more easily.

8.5.4. We appreciate *COBUILD*[2]'s attempt to give comprehensive information on language use based on The Bank of English. Now that *LDCE*[3] has greatly reduced pragmatic information, *COBUILD*[2] is the only major dictionary that provides learners with detailed information. We hope that editors of *COBUILD*[2] will improve, refine, and give further information on language use.   (H. Masuda)

---

1) *LDCE*[2] F12 Pragmatics and Dictionary.
2) See p. 185.
3) Leech (1983) explains this in terms of a COST-BENEFIT SCALE. See p. 107.

9. Summary

According to the backcover, *COBUILD*[2] covers over 75,000 references with 4,000 new words and meanings. The figure, like other dictionaries, seems to include all the information in bold face. The estimated main entries number only 33,752. The main area of new words comes from compounds, of which many are labeled as 'American.' Another area of new words comes from informal words. This might be the result of the improvement in the corpus that has focused considerable effort on the inclusion of the spoken languages.

COBUILD[2] has introduced *Superheadwords* to divide polysemous words. Although this is a welcome innovation, the policy for the introduction of *Superheadwords* is unclear and is inconsistent. Sometimes they are meaning-oriented, sometimes they are word-class-oriented. Sometimes words are unnecessarily divided under the different *Superheadwords*.

The adoption of *Frequency Bands* in the Extra column is another welcome innovation. If frequency indexes were put next to each distinct meaning of a word, they would be more helpful.

The prescriptive principle on which the suggested pronunciations are chosen characterizes the revision from the previous edition to the current one. Instead of describing how each word is pronounced as it is, using a complicated transcription system with superscript numerals as in the previous edition, the present edition attempts, whenever possible, to give a single pronunciation model for non-native learners to follow. This is a reasonable choice for a dictionary compiled for non-native learners of English, and this feature, together with the use of italicization for vowel reduction and the unique description of stress with underlining, makes this dictionary a good reference for RP pronunciation. Although GenAm pronunciation is included within the scope of the new edition, the description leaves some room for improvement, both in terms of accuracy (3.1) and the transcription system (3.2).

Thanks to the new layout, the sought for meaning has become much easier to find. It is especially helpful because all the meanings now start on a new line. The style of definitions in *COBUILD*[2] may be more informative than the one employed in most conventional dictionaries because

selectional restrictions and typical collocations of the entry words can be given in a natural sequence of words. Since definitions correspond to the grammatical labels in the Extra Column in most cases in the second edition, it has become more helpful than *COBUILD*[1].

However, it may be high time for all of us to think about both the good and bad aspects of defining sentences. The second edition can be criticized in two respects. (1) There seems to be no consistent rule for applying a variety of defining formulae to the specific case. (2) Many of *COBUILD*'s 'explanations' do NOT define the headword (especially in the noun entries). Instead, they explain the whole meaning of a phrase which consists of the headword and other elements. Some such phrases are treated as 'phrases' (hence, with a grammatical label 'PHRASES'), while other similar phrases have no label. We do not see any difference between them.

As far as the arrangement of senses is concerned, *COBUILD*[2] is heavily dependent on frequency as a guide to decide which of them should come first. Although *COBUILD*[2]'s frequency-based sense order should be appreciated, we believe it should not be applied too rigidly.

COBUILD[2] has defined and listed (almost) all the usage labels in the front matter. We welcome this. However, they are not 'labels' per se but explanatory notes. The style of presentation itself is acceptable, but the problem is that the dictionary is not consistent in this respect, either. Technical terms are described as such by means either of the label ('technical' with or without a specific field indication which is introduced by an 'in . . .' phrase) or of an explanatory note about the 'users' of the headword by means of a 'used by . . .' formula. The label 'spoken' seems to have been applied with no definite rule. It is good that the second edition marks British usage more often than the first edition.

COBUILD[2] is different from other similar learner's dictionaries in that it does not contain any 'usage notes', which are often used to explain semantic and other differences between synonymous lexemes.

Examples, reportedly all replaced by new ones, are often as difficult as those in the original edition, and there are examples the contexts of which are not clear. It is regrettable that the same criticism has to be repeated of the current edition.

It is also regrettable that there are more than "a few" entries where no example is given which could add 'something' to the information given in the definition.

There are three types of grammatical information in *COBUILD*[2]: the word class of the word, restrictions or extensions to its behavior, and the patterns in which the word most frequently occurs (p. xxiv). The word class is given in capital letters while the rest of the information is written mainly in small letters. The system of using the Extra Column to set out grammar notations is also employed in the new edition.

COBUILD[1] was the first dictionary to use the concept of ergativity and reciprocity in order to classify verbs. In *COBUILD*[2] the existence of ergative reciprocal verbs is recognized. It is noteworthy that the ergativity and reciprocity of phrases are also recognized in *COBUILD*[2]. The fact that phrases play an important part in English grammar is highlighted in the new edition.

COBUILD[2] uses a large number of codes and abbreviations in the grammatical notation. However, not all of them are self-explanatory. When the codes are not understandable at first sight, the users tend to give up deciphering the codes and eventually ignore the information conveyed by them. This should not be overlooked. It is expected that some of the grammatical information be explained in the definition and the use of hard-to-decipher codes be avoided in the next edition.

Pragmatics, a new feature to *COBUILD*[2], aims at giving information such as speaker's intention, emotion, emphasis, certainty and uncertainty, etc. to learners. Although some pieces of information on pragmatics were given in *COBUILD*[1] as part of its definition, and in other learner's dictionaries as well, information given in *COBUILD*[2] far exceeds that given in any of these other dictionaries. This information, we believe, plays an important role in the Dictionary because pragmatic strategies differ from language to language. Since it is a relatively new dictionary feature, inconsistency is unavoidable. Distinctions should be made in the defining sentences between the lexical meaning of the word and the utterance meaning. Editors should think very hard when deciding what sort of pragmatic information should most properly be given in the dictionary. Some sen-

tences have extra meanings that can be deduced from the general rule of pragmatics rather than from the particular word/phrase in the sentences. Shouldn't a dictionary give extra pages to explain them in those cases, instead of attributing 'meaning' to the word/phrase in question?

DICTIONARIES

CCEG   *Collins COBUILD English Grammar*. London and Glasgow: Collins ELT, 1990.

CIDE   *Cambridge International Dictionary of English*. Cambridge: CUP, 1995

COBUILD[1]   *Collins COBUILD English Language Dictionary*. London & Glasgow: Collins ELT, 1987.

COBUILD[2]   *Collins COBUILD English Dictionary*, New Edition. London: Harper-Collins, 1995.

CULD   *Chambers Universal Learners' Dictionary*. Edinburgh: W&R Chambers, 1980.

LDCE[1]   *Longman Dictionary of Contemporary English*. Harlow and London: Longman, 1978.

LDCE[2]   *Longman Dictionary of Contemporary English*, New Edition. Harlow: Longman, 1987.

LDCE[3]   *Longman Dictionary of Contemporary English*, 3rd ed. Harlow: Longman, 1995.

OALD[4]   *Oxford Advanced Learner's Dictionary of Current English*, 4th ed. OUP, 1989.

OALD[5]   *Oxford Advanced Learner's Dictionary of Current English*, 5th ed. OUP, 1995.

REFERENCES

Alexander, L. G. 1988. *Longman English grammar*. Harlow: Longman.

Burchfield, R. 1981. *The spoken word — A BBC guide*. London: BBC.

Francis, G. & Sinclair, J. 1994. '"I bet he drinks carling black label.": A Riposte to Owen on corpus grammar.' *Applied linguistics*, Vol. 15. No. 2.

Hausmann, F.J. and Gorbahn, A. 1989. 'COBUILD and LDOCE II: A comparative review.' *IJL*. Vol. 2. No. 1. pp. 44–56.

Higashi, N. et al. 1986. 'Gakushuyo amerika eigo jisho no hikaku to bunseki.' [A comparative analysis of the two dictionaries of American English — *LDAE* and *OSDAE* —] *Lexicon* No. 15. pp. 48–170.

Kojima, Y., et al. 1989. '*COBUILD* no bunseki.' [An analysis of *Collins COBUILD English Language Dictionary*]. *Lexicon* No. 18. pp. 39–154.

Kozaki, M. 1997. 'Heavy to Light no gogi ni tsuite: hiyu ni yoru imi no kakudai.' [On the meanings of Heavy and Light: metaphorical extensions of word meanings]. *Jissen Joshi Tandai Hyouron*. pp. 99–110.

Landau, S. I. 1984. *Dictionaries: The art and craft of lexicography*. New York: Scribner's.

Leech, G.N. 1983. *Principles of pragmatics*. Harlow: Longman.

Piotrowski, T. 1988. 'Review: *Collins COBUILD English Language Dictionary*.' *International review of applied linguistics*. Vol. 26. No. 3. pp. 253–55

Takebayshi, S. 1984. 'Amerika eigo no bird no boin' [Pronunciation of the American vowel in *bird*]. *Eigo Kyoiku*. Vol. 12. p. 67.

Trudgill, P. & Hannah, J. 1985. *International English*. 2nd ed. London: Edward Arnold.

Wells, J.C. 1982. *Accents of English*. Cambridge: CUP.

Zgusta, L. 1971. *Manual of lexicography*. The Hague: Mouton.