An Analysis of Collins COBUILD Advanced Dictionary of American English

HIDEO MASUDA  SATORU UCHIDA
MANAMI HIRAYAMA  AKIHIKO KAWAMURA
RUMI TAKAHASHI  YASUTAKE ISHII

1. Introduction

This paper is a critical review of Collins COBUILD Advanced Dictionary of American English (abbreviated as COBAm). COBAm, in the front matter, says “Through a collaborative initiative, Collins COBUILD and Thomson ELT, is copublishing a dynamic new line of learner’s dictionaries offering unparalleled pedagogy and learner resources” (Guide to Key Features). This statement explains how and why the dictionary was produced. It also says that the book was “specifically designed for the American market” (Introduction). We can say from these statements that COBAm is intended to be a new type of pedagogical dictionary that targets advanced learners of American English.

The aim of this paper is to examine whether COBAm really offers “unparalleled pedagogy and learner resources” for learners of American English in a learner friendly way. We have mainly used the paper version of COBAm for our analysis. However, special attention is paid to the CD-ROM version in Section 7, where the usability of the CD-ROM is intensively examined.

COBAm inherits most of the features of the COBUILD dictionaries. It is based on the Bank of English, a corpus that contains over 650 million words (according to the Introduction). Its definitions are written in full sentences. They give information on usage, register, typical context, and syntax as well as the meaning of the word. As in other COBUILD dictionaries, all the examples are chosen from the corpus “to show the collocates of a word — other words that are frequently used with the word we are defining — and the patterns in which it is used” (Introduction).

Since COBAm is based on the tradition of Collins COBUILD dictionaries, we have made comparisons with the previous editions, specifically with COB5, the latest edition of the British version. We have focused on words and phrases, pronunciations, senses, examples, and the information in the various boxes.

The major features incorporated in the dictionary, according to the Introduction, are DefinitionsPLUS (Collocations, Grammar, and Natural English) and Vocabulary Builders including Picture Dictionary, Word Web, Word Partnership, Word Link, Thesaurus, and Usage. Picture Dictionary boxes and Word Web boxes are colorfully illustrated to make the dictionary more attractive to the learners. The “Activity Guide” given in the front matter will help the learners master these features.

The following two changes may be worth mentioning: (1) COBAm has listed “a natural defining vocabulary” of “the 2,500 most common words of English”. No such list was given in COB5 although it mentions a natural defining vocabulary of the 2,500 commonest words of English. (2) The traditional Extra Column has been abolished, and the information in the Column such as grammatical pattern of verbs, pragmatics, and frequency bands is incorporated into the entry.

We have also compared COBAm with its rival dictionaries, namely LDOCE4 and OALD7 in Section 5 and LAAD2 in Section 3. This comparison with other learners’ dictionaries will clarify the characteristics of COBAm’s microstructure.

This paper is organized as follows. In Section 2, Uchida examines headwords, phrases and phrasal verbs to see how the dictionary has been Americanized from its source book, COB5. He also examines the super-headwords and Frequency Bands to see if there are any differences between these two books.

In Section 3, Hirayama discusses COBAm’s pronunciation, showing that the transcription is not always consistent and that this inconsistency can confuse learners. Pronunciation on the CD-ROM is also discussed.
In Section 4, Kawamura deals with COBAm's definitions, sense description, labels and pragmatics in comparison with those in COBS.

In Section 5, Masuda discusses the examples mainly focusing on the changes (deletions, additions, and alterations) from COBS, and the semantic contents of the examples. He examines what kinds of examples are appropriate for the learners.

In Section 6, Takahashi examines the nature and usefulness of Vocabulary Builders. She shows that the colorful Picture Dictionary boxes provide more of encyclopedic information rather than academic information. Word Webs, one of the properties of Vocabulary Builders, are innovative and provide useful encyclopedic information for the learners. In 6.6, Kawamura examines Usage notes and points out that some of the notes are too simple to be of much use to advanced learners.

In Section 7, Ishii deals with the CD-ROM versions of COBS and COBAm. He also compares them with the paper versions and discusses their respective advantages and disadvantages.

We have discussed some of the shortcomings as well as the valuable properties of COBAm. We hope our analysis will help improve the dictionary.

2. Headwords, phrases and phrasal verbs

This section discusses differences between COBAm and COBS, concerning headwords, phrases and phrasal verbs. The main purpose here is to examine how the dictionary is "Americanized" since COBAm is "specifically designed for the American market" (p. xii). After introducing the sampling methods in 2.1, the results are presented. Headwords and run-ons are discussed in 2.2, followed by discussion on phrases and phrasal verbs in 2.3. We will also see if there are any differences with regard to superheadwords in 2.4 and Frequency Bands in 2.5.

2.1. Sampling

In order to make a rough estimation of the newly added and deleted items between COBAm and COBS, two types of sampling methods are employed: (1) an examination of the two-page spread of every 100 pages of COBAm (bases to bass, ceasefire to centenary, credentials to crib death, duckling to dumby, fixed to flapper, handwriting to harass, intrigued to inverted commas, manner to margin, obtain to octave, pneumonia to pointing, religion to reminiscent, share capital to sheen, student to style, trumped-up to tuck, and wishy-washy to wizard, a total of 30 pages) and the corresponding pages in COBS, and (2) the whole pages of A (a total of 82 pages in COBAm, 85 pages in COBS) and L (which contains the center page of COBAm, a total of 53 pages in COBAm, 63 pages in COBS) of each dictionary. Method (1) is for a preliminary survey. The results are given in Table 2.1.

<table>
<thead>
<tr>
<th>Preliminary survey</th>
<th>Total</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headwords</td>
<td>606</td>
<td>+18</td>
</tr>
<tr>
<td>Run-ons</td>
<td>66</td>
<td>56</td>
</tr>
<tr>
<td>Phrases</td>
<td>60</td>
<td>33</td>
</tr>
<tr>
<td>Phrasal verbs</td>
<td>37</td>
<td>27</td>
</tr>
</tbody>
</table>

As is clear from Table 2.1, there are considerable changes between the two editions. Particularly, there are many headwords that are unique to each edition, and the numbers of run-ons, phrases and phrasal verbs in COBAm are remarkably fewer than COBS.

Given these promising results, a survey was conducted using method (2). The results will be discussed in the following sections together with those from the preliminary survey. The number of pages examined is 165 in total, which amounts to approximately 10% of COBAm.

2.2. Headwords and run-ons

In this subsection, we will compare COBAm and COBS in terms of headwords and run-ons. The following tables show the results:

---

Table 2.1 The results of the preliminary survey

As is clear from Table 2.1, there are considerable changes between the two editions. Particularly, there are many headwords that are unique to each edition, and the numbers of run-ons, phrases and phrasal verbs in COBAm are remarkably fewer than COBS.

Given these promising results, a survey was conducted using method (2). The results will be discussed in the following sections together with those from the preliminary survey. The number of pages examined is 165 in total, which amounts to approximately 10% of COBAm.
Table 2.2 Comparison of headwords between COBAm and COB5

<table>
<thead>
<tr>
<th>Headwords</th>
<th>Total Changes</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COB5</td>
<td>COBAm</td>
<td>Added</td>
</tr>
<tr>
<td>Preliminary survey</td>
<td>606</td>
<td>600</td>
<td>+18</td>
</tr>
<tr>
<td>Pages of A</td>
<td>1,828</td>
<td>1,807</td>
<td>+39</td>
</tr>
<tr>
<td>Pages of L</td>
<td>1,089</td>
<td>1,059</td>
<td>+36</td>
</tr>
<tr>
<td>Average per page</td>
<td>19.8</td>
<td>21.0</td>
<td>+0.6</td>
</tr>
<tr>
<td>Estimated total number</td>
<td>33,561</td>
<td>31,962</td>
<td>+913.2</td>
</tr>
</tbody>
</table>

Table 2.3 Comparison of run-ons between COBAm and COB5

<table>
<thead>
<tr>
<th>Run-ons</th>
<th>Total Changes</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COB5</td>
<td>COBAm</td>
<td>Added</td>
</tr>
<tr>
<td>Preliminary survey</td>
<td>66</td>
<td>56</td>
<td>0</td>
</tr>
<tr>
<td>Pages of A</td>
<td>172</td>
<td>125</td>
<td>+2</td>
</tr>
<tr>
<td>Pages of L</td>
<td>110</td>
<td>92</td>
<td>0</td>
</tr>
<tr>
<td>Average per page</td>
<td>2.0</td>
<td>1.7</td>
<td>+0.01</td>
</tr>
<tr>
<td>Estimated total number</td>
<td>3,390</td>
<td>2,587.4</td>
<td>+15.2</td>
</tr>
</tbody>
</table>

As shown in Tables 2.2 and 2.3, there are many items that are newly added as well as deleted items in COBAm. We will first examine the items newly adopted in COBAm, and then discuss the deleted items. We will also consider differences in the presentation of headwords between the two editions.

2.2.1. Newly added items

In regard to the sample pages examined, there are 93 items newly adopted in COBAm. The following are some random examples:

aboveboard, acclimate, all-you-can-eat, altostratus, alum, ameba, APB, arbor, area rug, credit limit, credit line, credit union, cremains, duds, DUI, land office, Lasik, Latina, laudanum, laugh lines, lawn bowling, lawn chair, letter carrier, lickety-split, listserv, log-rolling, manufactured home, podcast, student council, student loan, study hall

It should be noted that most of the newly adopted items are assigned AM (American) label. For example, duds, DUI, laugh lines, lawn bowling, lawn chair, letter carrier, lickety-split, listserv, log-rolling, manufactured home, and study hall are labeled as such. This is a natural consequence that follows from the fact that the current edition is modified for American English, and is surely one of the improvements made in COBAm in that it covers English vocabulary more extensively.

Other sources of new adoption are from sports, the Internet and computers. For example, one of the baseball terms line drive is added as well as some American football terms such as line man and line scrimmage. Concerning words related to the Internet and computers, there are netspeak terms such as lol (laughing out loud) and the names of the Internet services such as podcast. This change may be an attempt to reflect the cultural differences between the United States and Britain, and the expanding vocabularies in the domain of computers.

With regard to run-ons, there are few additions to the American edition including acclimation and airsickness. On the other hand, as will be discussed in the section to follow, there are many run-ons that only appear in COB5.

2.2.2. Deleted items

Let us now turn our attention to the deleted items in the American edition. Some random examples are given below:

abstruse, access course, acid house, active service, actual bodily harm, adduce, air-con, air hostess, airing cupboard, air pistol, airport novel, airy-fairy, lace-ups, lateral thinking, lavatorial, lavatory paper, leant, learnt, lie-down, lino, lippy, listed, lock-keeper, locum, lollipop, lolly, long-hours culture, long-wearing, loo, manor, manor house, manservant, marge, obtrude, och, trunk road, TUC, with-profits, witness box, witter

There are 150 items that are deleted in the sample pages of COBAm. It should be noted that most of these items are assigned BRIT (British) label in COB5. The examples are lace-ups, lavatory paper, leant, learnt, lie-down, manor, manservant, trunk road, witness box and witter.
Headwords with other labels such as [irish] (e.g. och) and [mainly BRIT] (e.g. lateral thinking) are also excluded from the present edition in many cases. This may be reasonable given that the present edition is aimed to describe American English. However, if the user comes across text written in British English and finds items that s/he wants to look up in the dictionary, it is inconvenient that the user never has a chance to access these items in COBAm. One solution would be to retain these items in the CD-ROM edition where space is not an obstacle, but they do not appear in the present CD-ROM version either.

With regard to run-ons, there are also many items that are not included in COBAm but are included in COB5. In the sample pages examined, there are 77 such items. Some of them are listed below:

- abjectly, abominably, abstractly, abuser, acclimatization, acoustically, acrimoniously, activation, affirmatively, affordability, agonizingly, alphabetically, ambivalence, analytically, anciently, astoundingly, attractively, dumbing, dumbly, flagrantly, hanging, laterally, legalization, leveraged, lightly, loading, lobbying, loudness, manoeuvring, obtuseness, occupationally

The number of deleted items in the American edition is much larger than those added; the estimated number of deletions is 761, which is more than 50 times that of the estimation of newly added run-ons. The reason for this drastic reduction is presumably the fact that more space is needed for the pictures and columns newly introduced in COBAm (see Section 6 for information on the new features of COBAm).

### 2.2.3. Presentation of headwords

As a result of the modification to the American English version, there are some changes regarding the presentation of headwords. Three points are to be noted here. First, COBAm employs American spelling. For example, labour is changed to labor in the American edition, which at the same time means that some compounds such as labour camp and labour force are modified to labor camp and labor force respectively. Some other examples are lacklustre to lackluster, left-of-centre to left-of-center, appall to appal, and leveller to leveler. This kind of change is also observed in the definitions of COBAm (see Section 4 for detail).

Second, some changes are made to the hyphenations in compounds. For example, hyphens are deleted in anti-hero, lamp-post, life-saving, line-up and link-up, and they appear in COBAm as antihero, lamp-post, lifesaving, lineup and linkup. The opposite change is also found, i.e. spaces are replaced with hyphens as in air conditioned to air-conditioned and air conditioning to air-conditioning.

Lastly, changes are also made concerning the use of period. For instance, aka, appt and Lt in COB5 are changed to a.k.a., appt. and Lt respectively. These presentational changes may be the result of more extensive use of corpora, and this can be seen as an improvement to reflect the status quo of American English, although it is regrettable that alternative spellings are not provided in many cases.

### 2.3. Phrases and phrasal verbs

Let us now move on to phrases and phrasal verbs. The following tables display results from the examinations of the sample pages:

#### Table 2.4 Comparison of phrases between COBAm and COB5

<table>
<thead>
<tr>
<th>Phrases</th>
<th>Total Changes</th>
<th>COB5</th>
<th>COBAm</th>
<th>Added</th>
<th>Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preliminary survey</td>
<td>+1</td>
<td>60</td>
<td>33</td>
<td></td>
<td>-28</td>
</tr>
<tr>
<td>Pages of A</td>
<td>0</td>
<td>228</td>
<td>182</td>
<td></td>
<td>-46</td>
</tr>
<tr>
<td>Pages of L</td>
<td>+1</td>
<td>175</td>
<td>99</td>
<td></td>
<td>-77</td>
</tr>
<tr>
<td>Average per page</td>
<td>4.01</td>
<td>2.6</td>
<td>1.9</td>
<td></td>
<td>-0.9</td>
</tr>
<tr>
<td>Estimated total number</td>
<td>+15.2</td>
<td>4,407</td>
<td>2,891.8</td>
<td></td>
<td>-1,369.8</td>
</tr>
</tbody>
</table>

#### Table 2.5 Comparison of phrasal verbs between COBAm and COB5

<table>
<thead>
<tr>
<th>Phrasal verbs</th>
<th>Total Changes</th>
<th>COB5</th>
<th>COBAm</th>
<th>Added</th>
<th>Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preliminary survey</td>
<td>0</td>
<td>37</td>
<td>27</td>
<td></td>
<td>-10</td>
</tr>
<tr>
<td>Pages of A</td>
<td>0</td>
<td>24</td>
<td>18</td>
<td></td>
<td>-6</td>
</tr>
<tr>
<td>Pages of L</td>
<td>+2</td>
<td>133</td>
<td>104</td>
<td></td>
<td>-32</td>
</tr>
<tr>
<td>Average per page</td>
<td>+0.02</td>
<td>1.1</td>
<td>0.9</td>
<td></td>
<td>-0.3</td>
</tr>
<tr>
<td>Estimated total number</td>
<td>+30.4</td>
<td>1,864.5</td>
<td>1,369.8</td>
<td></td>
<td>-456.6</td>
</tr>
</tbody>
</table>
What is newly adopted in COBAm are cross the line (step over the line) for phrases, and load down, look in and luck out for phrasal verbs. There are also many deleted items for both phrases and phrasal verbs. Some random examples are given below:

[Deleted phrases]
with one accord, as against, comes of age, under age, in aid of, up in the air, alive and kicking, fly the flag, by and large, late in the day, least of all, at length, to the letter, come to life, in someone else's pockets

[Deleted phrasal verbs]
ask around, creep up on, flag down, hang around, land up, lay in, let down, let on, lift off, listen out, live through, live together, look ahead, be lumbered with, lust after

One remarkable change is that the American edition seems to cut out the section of “phrases” columns, which list a number of phrases in the British edition. As a result, for example, the phrases such as least of all, not least and to say the least at the entry of least, and get a life, in all my life and larger than life at life are excluded from COBAm. There are also some cases where items remain but the number of definitions is decreased. To take light up, for example, there are three definitions with the label PREP-PHRASE in COB5 but the third definition is not found in COBAm, namely “If you light up, you make a cigarette, cigar, or pipe start burning and you start smoking it”. Again, this deletion may be to save space for the pictures and new columns, but such deletions may have reduced the necessary information for a foreign learner.

It should not be overlooked that some of the deleted phrases and phrasal verbs still remain in the definitions and examples in COBAm, a fact which causes a serious problem. Consider the following examples:

**survival** 2 If you refer to the survival of a person or living thing, you mean that they live through a dangerous situation in which it was possible that they might die.

**bloom** 4 Not many economies bloomed in 1990, least of all gold exporters like Australia.

An analysis of Collins COBUILD Advanced Dictionary of American English

The former is taken from the definition of survival and the latter is from the examples of bloom. The definition and the example above contain phrases live through and least of all which are excluded from COBAm. This is a problematic situation since the definitions of these phrases are not available in COBAm in spite of the fact that they are employed in the definitions and examples. This means that if the user cannot figure out the sentences with such phrases, s/he has to consult other lexicographic resources. This should be seen as a serious defect of COBAm.

A word should be said that the label PREP-PHRASE seems to be excluded from the present edition. There are 29 items labeled as such in the sampling pages of COB5, but 26 of them are changed into PHRASE and one is deleted (in advance of). As for the other two, under the aegis of and à la, they are still labeled as PREP-PHRASE. This might be an editorial mistake since most of them seem to be automatically changed to PHRASE. This change in COBAm makes the descriptions simpler because there are some cases where the distinction between PREP-PHRASE and PHRASE is not clear in COB5. For example, in aid of is treated as a phrase, while in the light of is as a prep-phrase in COB5, but both of them are treated as phrases in COBAm.

Last but not least, it should be pointed out that as a result of the deletion of the Extra Column, which is counted as one of the most remarkable changes between the two editions, it becomes very difficult to search phrases and phrasal verbs in COBAm. In COB5, the user can instantly locate the phrases and phrasal verbs of the target word by simply searching down the Extra Column. However, the American edition requires the user to read through the whole entry to find the phrases and phrasal verbs s/he wants to look up. Thus, the user may also find it difficult to search labels assigned to the definitions (see Section 4). Again, this may be closely related to the space issue, but one may wonder if such a change is worth the cost of abandoning one of its distinctive features.

### 2.4. Superheadwords

In this subsection, we focus on the changes concerning the superheadwords. Superheadwords are headwords where signposts are presented that
are intended to help quickly find the definitions and capture the picture of the meanings of polysemous words. In the sample pages of COBAm, newly added items are found at all, leave and line and no deletion is made in the present edition.

There are two types of shortcuts listed at the superheadwords. For words such as lay, like, line and live, the shortcuts indicate the word classes (e.g. noun use, verb use, etc.). The other type is to present the synopses of the meanings, for instance, at the entries of lead and light (i.e. 1 BEING AHEAD OR TAKING SOMEONE SOMEWHERE, 2 SUBSTANCES FOR lead, and 1 BRIGHTNESS OR ILLUMINATION, 2 NOT GREAT IN WEIGHT, AMOUNT, OR INTENSITY, 3 UNIMPORTANT OR NOT SERIOUS FOR light). However, there are cases where the shortcuts are somewhat confusing for the user in that both word class labels and meaning labels are listed at the same place. Consider the following example:

<table>
<thead>
<tr>
<th>long</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 TIME</td>
</tr>
<tr>
<td>2 DISTANCE AND SIZE</td>
</tr>
<tr>
<td>3 PHRASES</td>
</tr>
<tr>
<td>4 VERB USES</td>
</tr>
</tbody>
</table>

This is not a straightforward way of listing shortcuts, since the first two lines show the meaning of the headword, whereas the latter two indicate the usages (cf. Masuda et al. 2003: 10).

It should also be pointed out that the policy is inconsistent. For example, the superheadword at look only refers to the meanings (e.g. 1 USING YOUR EYES OR YOUR MIND, 2 APPEARANCE), but not to the phrasal verbs regardless of the fact that there are nearly twenty of them. In contrast, the shortcuts for the phrases or phrasal verbs are listed at the superheadwords of leave, line, long, etc. Besides, as is pointed out in Masuda et al. (2003: 11), there are some cases where completely different words are listed at the same place such as Miss (1 USED AS A TITLE OR A FORM OF ADDRESS) and miss (2 VERB AND NOUN USES). Overall, the policy of the superheadwords leaves much to be improved.

2.5. Frequency Bands

Frequency Bands are one of the distinctive features of COBUILD dictionaries. They indicate how frequent the headwords are used based on the huge corpus called the Bank of English “which now contains over 650 million words of contemporary English” (p. xii). This information is useful not only for students but also for teachers in deciding whether to emphasize the words they teach. These bands had five grades in COB3, but they are reduced to three in COB4 and COB5, and COBAm also employs the three degrees of bands.

A question arises as to whether Frequency Bands reflect the differences of frequency between American and British English. In fact, some modifications are made as shown below. Numbers in parentheses indicate the grade:

- [The item whose Frequency Band is deleted]
  accommodation (1)

- [Items with Frequency Bands in COBAm but not in COB5]
  anymore (2), April (3), Asian (2), audio (2), August (3), Avenue (2), lawsuit (2)

As for the deleted item, accommodation, the deletion of Frequency Bands can be definitely justified because it is labeled as [BRIT]. Also, the addition of Frequency Bands at anymore (the corresponding form of British English is “any more”) and lawsuit can be defended for the opposite reason that these two items are more extensively used in the United States. As for the other newly added items, most of them are calendaric terms such as April and August, and geographical names such as Asian. This may be due to the change of the policy as to the use of Frequency Bands for these kinds of items. One thing to be noted here is that the policy for the adoption of Frequency Bands at calendaric and geographical names is unstable through the editions of COBUILD dictionaries: Frequency Bands for these items appear in COB3, but disappear in COB4 and COB5, and they are reintegrated in COBAm. As is pointed out in Masuda et al. (2003: 15), it is questionable that Frequency Bands at these words are of great help for the user especially when considering that
the dictionary is aimed at the advanced level: they are such basic words that the user of the dictionary is expected to already know them. Unfortunately, there is no explanatory section of Frequency Bands in COBAm, so the reason for the reintegration remains unclear. (Section 2 by Uchida)

3. Pronunciation*

In this section, COBAm's pronunciation is discussed, specifically focusing on whether it reflects the actual situation of American English. COBAm states that its transcription is not based on American English alone; rather, American English is one of the (two) varieties on which their pronunciations are based. For example, a systematic bilectal notation is employed for rhoticity (3.1.1, 3.2.1), but American forms are often the only pronunciation given in the entries (3.2 (paper version), 3.3 (CD-ROM)). Furthermore, some features described in the literature as changes in progress that are not discussed in the paper version of COBAm are present on the CD-ROM (3.3). General discussion is given in 3.1. Minor typos are pointed out in 3.4.

Two notes are in order. First, following Masuda et al. (1997, 2003, 2005), we call the pronunciation section of the introductory part of the dictionary the Guide. Second, stress is omitted unless otherwise noted.

3.1. General characteristics of transcription

This section discusses general properties of COBAm transcription, focusing on dialectal variation (3.1.1, 3.1.2), levels of transcription (3.1.3) and transcription symbols (3.1.4).

3.1.1. Dialect(s)

The philosophy of COBAm transcription is given in the Guide as follows: "The basic principle underlying the suggested pronunciations is 'If you pronounce it like this, most people will understand you'" (p. xvi). This same statement appears in COB5 (p. xxiii) and the two previous editions of COB, as reviewed in Lexicon (new edition in Masuda et al. (1997), COB3 in Masuda et al. (2003)). What follows this statement in the Guide is discussion on the dialect(s) in COBAm transcription. While one may expect it to be based on American variety(s), it is not fully so; the Guide mentions British English as well as American English as the basis for their pronunciations: "The pronunciations are therefore broadly based on the two most widely taught accents of English, GenAm or General American for American English, and RP or Received Pronunciation for British English" (p. xvi). It goes on to say how they treat the two dialects in the entries: "For the majority of words, a single pronunciation is given, as most differences between American and British pronunciation are systematic. Where more than one pronunciation is common and the difference is not accounted for in the notes below, alternative pronunciations are given" (p. xvi). This statement might be interpreted thusly: first, when GenAm and RP pronunciations differ systematically, it is not clear which dialect form is given in the entries, since it is not specified (though one may expect it to be a GenAm form since it is an American English dictionary). Second, there are four "notes", three of which discuss differences between GenAm and RP (ask words, rhoticity and /ou/). These three features clarify which form, GenAm or RP, is given in a particular entry. Third, when the pronunciations differ between GenAm and RP but not systematically, both GenAm and RP forms would be given. The Guide does not discuss the subject any further.

On the first point, as discussed below in 3.2, as far as the selected GenAm diagnostic features are concerned, COBAm gives only the GenAm forms in most cases. On the second point, in the notes, the difference between GenAm and RP is explained for each feature, but except for ask words, it does not specify which dialect will be given in the entries; for ask words, it is specified that GenAm forms are given. Even so, the forms are not given consistently. For example, although ask words are in most cases assigned to the GenAm vowel, there are several words which are assigned to both GenAm and RP vowels (see 3.2.4 for the details). In another note, on rhoticity, COBAm appears to systematically show both GenAm and RP forms in the entries, even though the difference is already "accounted for" in the notes of the Guide. A bilectal transcription is furthermore confusing for the users, who would expect American English pronunciation as it is an American English dictionary. We will return to this point in
detail in 3.2.1. On the third point, as discussed in 3.2.8, for the great majority of lexical items whose pronunciations differ non-systematically between GenAm and RP, both forms are not always given.

To conclude, the statements in the Guide do not always make clear what dialect(s) are shown in the entries. Furthermore, dictionary users may be confused by the appearance of RP as well as GenAm in a dictionary whose title may lead them to expect only American pronunciations. Compare this with LAAD2 and COB5. LAAD2's focus is on American pronunciations, as stated in the Guide: “This dictionary shows pronunciations used by most American speakers” (p. ii). COB5 is bilectal, but it is not confusing since the bilectal notations are explained in the Guide and the users are not likely to expect otherwise, unlike COBAm.

3.1.2. Variation

GenAm has variation, as the definition of the term in the literature indicates: “This [General American] is a convenient name for the range of United States accents that have neither an eastern nor a southern colouring” (Wells 1982: 10). Variation within American English is not discussed in COBAm; it is not entirely clear from the Guide whether the alternative pronunciations in the entries are meant to reflect variations within American English or differences between GenAm and RP. The Guide seems to suggest that it is the latter, i.e. the variants are to indicate the differences between GenAm and RP, not differences within GenAm; the Guide makes no explicit mention of variation within GenAm and clearly states that COBAm is based on both RP and GenAm, not GenAm only.

This can be compared with another American English dictionary, LAAD2; LAAD2 discusses variation in the Guide, and the variants in the entries are said to be “possible and...commonly used by educated speakers” (p. ii). Even when the alternative pronunciations are not given in the entries, it is recognized that variation exists; the Guide gives examples of variation, such as the cot-caught merger and Yod dropping (discussed below in 3.2.3 and 3.2.5, respectively), and acknowledges further variation that is not discussed: “We have not, however, shown all possible American pronunciations” (p. ii).

This difference between the two dictionaries probably stems from the philosophy on transcription. COBAm’s aim is to provide variant(s) that “most people will understand” (p. xvi); thus, discussion of variation might not have been a high priority.

3.1.3. Levels of transcription

There is a difference between COBAm/COB5 and LAAD2 with respect to levels of transcription. COBAm and COB5 employ more underlying-oriented transcription, with very few allophonic features, while LAAD2 provides more surface-oriented transcription, with some allophonic features that are characteristic of American English.

In COBAm and COB5, there is virtually only one allophonic feature, namely, syllabic consonants ([l] and [n]). In LAAD2, other allophonic features are also shown in the entries, including r-colored “s”, tap for /t/ (shown as “Atual”), glottal stop replacing or accompanying /t/ (“t’”) and optional deletion of /t/ and /d/ (shown in italic “r”, “d”), in addition to the syllabic consonants. These, and an additional feature, epenthetic [t] between a nasal and a fricative in the sequence of /n/ /t/ /d/ are described in the Guide (pp. i—iii), among which tap, glottal stop, t/d deletion and epenthetic [t] in /n/ /t/ /d/ are moreover highlighted, under “American Pronunciation” (p. i).

3.1.4. Symbols

COBAm’s transcription symbols are claimed to be basically IPA: “The symbols used in the dictionary are adapted from those of the International Phonetic Alphabet (IPA), as standardized in the English Pronouncing Dictionary by Daniel Jones (14th Edition, revised by AC Gimson and SM Ramsaran 1988)” (p. xvii). This is exactly the same as COB5, except COB5 has an additional phrase after this: “. . ., for representing RP” (p. xvi). LAAD2 also basically follows IPA (p. ii).

There are, however, a couple of symbols in COBAm that are not from the IPA as described in the Phonetic Symbol Guide (Pullum and Ladusaw 1996) or found in EPD14. These are /y/ (j) in IPA and EPD14 and the use of underline for marking stress (vertical stroke in both current IPA
and EPD14). These deviations are not explained in COBAm, nor in COB5. LAAD2 also deviates from the IPA in some symbols, but these deviations are discussed in the Guide.

3.2. Diagnostics for American English

This section examines how COBAm transcribes American pronunciation in the entries, focusing on several diagnostic characteristics. We focus on the paper version of the dictionary in this section, and discuss the sounds on the CD-ROM later (3.3). We will show that while COBAm's transcription is by and large only General American (GenAm), in some cases, specifically the contrasts in the low-back vowels (3.2.2) and the bilectal transcription in rhoticity (3.2.1), COBAm's transcription may be unexpected, based on the literature on American English pronunciation.4


In the following subsections, we discuss the selected features of GenAm which include rhoticity (3.2.1), vowel contrasts in the low back area (3.2.2, 3.2.3), ask words (3.2.4), Yod dropping (3.2.5), status of /h/ in /hw/ onset cluster (3.2.6), vowel merger as well as neutralization before /r/ (3.2.7) and lexical items (3.2.8).

3.2.1. Rhoticity

Rhoticity, or r-full/-lessness, is one feature that divides GenAm from RP (e.g. Ladefoged 2001: 28, Wells 1982: 75–76); GenAm is rhotic, with [r] occurring before consonants and word-finally as well as in other contexts (or, arguably, [r] being present in the coda), while RP is non-rhotic, with [r] occurring only before a vowel. In COBAm, the Guide mentions this difference, and the coda /r/ is shown in the entries as expected for GenAm. However, COBAm is best interpreted as bilectal in showing r-full/-lessness. We discuss this below.

There is a schwa /a/ before /r/ in the entries (at least at the end of a word, thus before a tautosyllabic /r/) unless the preceding vowel is low back (a, a) or central (a, a), as shown in (1).

(1) /a/ before /r/ after a vowel that is not low back or central (boxes are ours)

<table>
<thead>
<tr>
<th>Symbols in the entries</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>a&lt;sub&gt;r&lt;/sub&gt;</td>
<td>beard, near</td>
</tr>
<tr>
<td>a&lt;sub&gt;r&lt;/sub&gt;</td>
<td>care, fair</td>
</tr>
<tr>
<td>u&lt;sub&gt;r&lt;/sub&gt;</td>
<td>cure, poor, pure, tour</td>
</tr>
<tr>
<td>a&lt;sub&gt;r&lt;/sub&gt;</td>
<td>flour, flower, sour, tower</td>
</tr>
<tr>
<td>a&lt;sub&gt;r&lt;/sub&gt;</td>
<td>fire, tire</td>
</tr>
</tbody>
</table>

There is no /a/ when the vowel is low back or central, as shown in (2).

(2) No /a/ before /r/ if the preceding vowel is low back (a, a) or central (a, a)

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>a&lt;sub&gt;r&lt;/sub&gt;</td>
<td>car, cart, far, heart, tar</td>
</tr>
<tr>
<td>o&lt;sub&gt;r&lt;/sub&gt;</td>
<td>cord, more</td>
</tr>
<tr>
<td>a&lt;sub&gt;r&lt;/sub&gt;</td>
<td>third, turn</td>
</tr>
<tr>
<td>o&lt;sub&gt;r&lt;/sub&gt;</td>
<td>forgive, forgotten, teacher</td>
</tr>
</tbody>
</table>

We are not certain whether there is any phonological or phonetic motivation to divide into these two groups, i.e. evidence for which one group of words has /a/ before /r/ (ar, ar, etc.) and the other group does not (ar, ar, etc.). There are two possibilities, but neither of them is likely. One is that /a/ is syllabic on its own; thus, for example, near is disyllabic [ni.ar] (period indicates syllable boundary), not monosyllabic. However, this is clearly not the case judging from how the stress is marked: stress is marked in COBAm by “underlining the vowel in the stressed syllable” (p. xviii), and in near, both /i/ and /a/ are underlined, /ni.ar/. If it is disyllabic, this indicates that both the first and second syllables are stressed. This would be odd, given its normal pronunciation; if this word is disyllabic, only /i/ should be underlined, /ni.ar/, with an unstressed second
syllable. The other possible explanation for this transcription is to indicate an inglide character of [a] before [a] (Wells 1982: 153 for near and words with the same vowel). However, it is unlikely that this transcription is intended to represent such fine phonetic detail, given the level of phonetic detail that COBAm provides elsewhere (see 3.1.3 for the transcription level).

The pronunciation table is not revealing either: the table in the Guide does not include Vr sequences; the table on the inside of the back cover does not have schwa at all, and gives [i]/[r]/, [e]/[r]/, etc. instead. Note that this notation is not employed in the entries.

Thus, the notation of [a] before the coda [r] is most likely to have been employed to incorporate GenAm and RP pronunciations into one transcription. However, this interpretation is not user-friendly for two reasons. First, one does not expect to see RP pronunciation systematically in an American English dictionary. Second, there is no explicit explanation of how to interpret this notation; thus, users of the dictionary need to learn it by themselves, presumably from the explanation of the distribution of coda [r] between GenAm and RP given in the Guide (p. xviii). This can be compared with COBS, a bilectal dictionary in its policy, which employs a similar notation to COBAm but explicitly states in the Guide that schwa is not present in GenAm: “Some of the complex vowel sounds shown in the table above are simplified in GenAm. The vowel sound in ‘fire’ is shown as /a/ia/. This represents the pronunciation ‘fire’ in RP, but in GenAm the pronunciation is not /a/air/, but /a/air/”. So ‘fire’, ‘flour’, ... are pronounced /faiə/, /flauər/, ... in RP, but /fiər/, /flauər/, ... in GenAm” (p. xxiii). In COBAm, all but the last sentence is omitted in the Guide; this omission makes the COBAm notation hard to interpret properly.

3.2.2. Contrast in the low back area

Some English dialects have three contrasting vowels in the low back area. For example, in RP, the (stressed) vowel in calm, father, palm contrasts with the vowel in hot, lot, odd, further contrasting with the vowel in caught, ought, sauce. In GenAm, due to a merger, the contrast is generally assumed to be a two-way one, with unrounded (calm, palm, lot, odd), transcribed broadly [a], against rounded (caught, cloth, sauce), transcribed broadly [o] (e.g. Labov et al. 2006: 13–14, Ladefoged 2001: 29, Wells 1982: 120–124).

In COBAm, somewhat unexpectedly, a three-way contrast is employed. This is indicated in the list of vowels in the Guide and also in the list given inside of the back cover. We checked at the entries the words listed in Wells (1982) under LOT words (p. 131), CLOTH words (pp. 136–137, only (a), since the rest is noted as varying considerably in the United States (p. 136)), THOUGHT words (pp. 145–146) and PALM words (p. 144, only (a), i.e. native words). LOT words are invariably assigned to [a] in COBAm, differentiated from PALM words, which are again invariably transcribed [a], further different from CLOTH and THOUGHT words, which are all [3], thus a three-way contrast [a] vs. [a] vs. [3], as summarized in Table 3.1.56

| Table 3.1 Low back vowels in COBAm, COBS and LAAD2 |
|------------------|------------------|------------------|------------------|
|                  | CLOTH/THOUGHT    | LOT              | PALM             |
| COBAm            | [a]              | [a]              | [a]              |
| COBS             | [a]              | [a] (= a:)       | [a]              |
| LAAD2            | [a]              | [a] (= a:)       | [a]              |

Notes: (1) CLOTH/THOUGHT: In COBS, AM pronunciation is taken. In COBS, doss is assigned to [s]. In LAAD2, Australia, Austria, Austin and gone have [a] as an alternative; (2) PALM: In LAAD2, water has [s] as an alternative.

COBS and LAAD2, on the other hand, employ a two-way contrast, as expected from the literature, and as shown in Table 3.1. For these two dictionaries, among the words checked for COBAm, a subset was sampled (the words that define the categories (i.e. cloth, lot, palm, thought) and the first two words in each line in the above-mentioned list in Wells (1982)). COBS uses three symbols: [a], [o] and [o:]. However, there is a note in the Guide that [o] is the same as [o:] in GenAm. COBS employs a two-way contrast in this way. In LAAD2, the two-way contrast is expressed explicitly by using two symbols, [a] for LOT and PALM words, distinguished from [s] for CLOTH and THOUGHT words. In sum, the generally assumed GenAm two-way contrast is most explicitly shown in
LAAD2, less explicit but employed in COB5, and not employed in COBAm.

3.2.3. Further merger in the low back area — cot–caught merger

The literature reports an in-progress GenAm merger between the vowels discussed in the previous section, i.e. merger between /a(1)/ (LOT/PALM) and /a(2)/ (CLOTH/THOUGHT). As a result, pairs such as caller–collar, cot–caught, don–dawn, knotty–naughty, stock–stalk are all homophones. Wells (1982: 473-475) calls this the THOUGHT–LOT Merger. Ladefoged (2001: 28) notes this incomplete merger for “[m]any Midwestern speakers and most Far Western Speakers”. From Labov et al. (2006), the western half of the United States has the merger; the isogloss outlining the areas of merger (both in perception and production) includes Canada, the West, Eastern New England and western Pennsylvania (p. 59, referring to Map 9.1). Labov et al. also describe a real-time expansion of the merger by comparing their results with those in the past surveys (PEAS and Labov’s 1966 survey).

COBAm and COB5 do not describe the merger, as was discussed in 3.2.2. In LAAD2, the merger is not found in the sample words or in the list of vowels in the Guide, but it is discussed in the Guide (p. ii) as a feature of “many speakers”. We will return to this merger later (3.3) in the discussion of the CD-ROM.

3.2.4. ask words

Wells (1982: 133-135) indicates that a set of words, represented by bath (hereafter ask words), have stressed /æ/ in GenAm and /a/ in RP. The Guide of COBAm has a note on this dialectal difference, giving the phonological environments for the distribution of /a/ in RP. This environment is not given in COB5, which also has a note on ask words in the Guide.

The words listed in Wells as BATH words are all expected to have /æ/ in COBAm, since the note in the Guide states that it shows “only the GenAm version” for this (p. xviii). In order to see what vowel(s) are assigned to ask words at the entries, we checked the words listed in Wells

(1982: 135). The distribution in COBAm is given in (3). As expected, the great majority of the words are transcribed only as /æ/ , as shown in (3c).

(3) Transcription of bath words in COBAm
a. One has /a/ only: Slav
b. Some show variation between /æ/ and /a/:
   i. /a, æ/: advantage, aghast, ask, bask, basket, glass, hasp, Iran, lather, pass, plant
   ii. /æ, a/: aunt, Iraq, Sudan

c. The rest — the great majority — have only /æ/: bath, can’t, class, command, example, graph, half, laugh, ranch, etc.

We checked a subset of the words in LAAD2, namely, the word bath and the first two words in each line in Wells’s list. In LAAD2 also, the great majority of words is assigned exclusively to /æ/.

There are several words that COBAm gives an alternative /a/ for in addition to the expected /æ/, as shown in (3b). Note that some are everyday words such as ask. As discussed in 3.1.2, this alternative is likely to introduce an RP form, not another variant within GenAm. LAAD2 has only one instance of this type of word, as far as the words sampled for LAAD2 are concerned, which is Iraq. In both COBAm and LAAD2, Slav is exclusively assigned to /a/ (3a).7

In sum, COBAm and LAAD2 both provide the American form, /æ/. COBAm moreover records more variation than LAAD2, although the additional forms may be meant to be RP rather than GenAm.

3.2.5. Yod dropping

The literature (e.g. Wells 1982: 247-248) describes elimination of /j/ from historical /ju/ (or from /jio/ when /r/ follows), called yod (/j/) dropping. Dropping yod “after palatals, /z/, and clusters with /l/” (Wells 1982: 247) is common in both GenAm and RP, but generalizing the context to after all coronals including /t, d, n, θ, s, z, l/ is a tendency of GenAm, especially in stressed syllables. For example, Labov et al. (2006), investigating the minimal pair dew–do, shows that the contrast is predominantly not maintained, the distinction being “found today only in a minority of speakers in the South, concentrated in central North Carolina and the
lower Gulf States” (p. 55, in Map 8.3). In LPD2, in *deek, student*, and *Tuesday*, pronunciation without /j/ is provided as primary for GenAm, though forms with /j/ are also listed as less common variants. Moreover, in *student*, the result of an opinion poll carried out in 1993 by Shitara shows a large preference for the absence of /j/ (88%). In unstressed (or “weak” in Wells’s term) syllables, Wells indicates that the dropping is less widespread. Instead, he introduces the tendency toward another process, namely, coalescence of /j/ with the preceding consonant. He also observes a similar co-articulation in /nj/ and /lj/ sequences, i.e. [n] or [l] and [x] or [j], respectively (p. 248).

In *COBAm, COB5* and *LAAD2*, we checked all the words appearing in Wells (1982: 247–248) to see if yod is dropped.9) In primary stressed syllables (e.g. *duke*, *new*, *resume*, etc.) all three dictionaries show the tendency described in the literature, i.e. toward the absence of yod (e.g. /duk/, /nul/, *re/zoom*, etc.). In unstressed syllables (annual, educate, failure, issue, situate), yod is present in all three dictionaries, which also agrees with the literature. With respect to the other process, i.e. yod coalescence, it occurs if the coalesced, or fused, consonant is included in the English phonemic inventory (e/d3/ucate, i/j/ue, si/tj/uate), and does not when they are not contrastive (a/ny/ual, fai/ly/re). This is probably expected, since dictionaries normally do not transcribe the allophonic details to this extent. In *COB5, issue* has an alternative form not undergoing the coalescence i/sj/ue, as well as the coalesced form; this is not expected from the literature.

**Attitude, avenue** and **residue**, which are described as having unstressed syllables without vowel reduction in *COBAm* and *COB5* and having syllables with secondary stress in *LAAD2*, invariably lack yod in *COB5* and *LAAD2*, which conforms to what Wells (1982: 247) describes. In *COBAm*, however, it depends on the word: for **avenue** and **residue**, both forms — with and without yod — are given, while for **attitude**, yod is invariably dropped.

To summarize, *COBAm* gives the American pronunciation in primary stressed syllables. In unstressed syllables without vowel reduction, two of three items deviate from Wells’s description in that a form with yod is given (in addition to the form without yod). As has been discussed before, not all variants may represent GenAm in *COBAm* and these additional forms may show RP variants. In other unstressed syllables, yod coalescence is expected and found in *COBAm*.

### 3.2.6. /h/ before /w/ in the onset

Wells (1982) describes, as a change in progress, loss of /h/ in the onset /hw/ cluster in words such as *whale, wheat, whine*: “In North America, /hw/ is still a widespread usage, . . . . But Glide Cluster Reduction [= loss of /h/ in /hw/ cluster] is clearly on the increase, particularly in large cities” (pp. 229–230). Labov et al. (2006) finds this to have been widespread in the GenAm areas: “In the middle of the twentieth century, the distinction between /hw/ and /w/ in *whale* vs. *wail*, *which* vs. *witch*, etc. was maintained by most American speakers, with the exception of southern Maine; Boston; the Mid-Atlantic area, including Hudson Valley; and the Savannah–Charleston coastal region. In the ANAE data [= Labov et al.’s (2006) survey], the distinction is made only by a scattering of speakers, mainly concentrated in the Southern states” (p. 50, in Map 8.1).

In *COBAm* and *LAAD2*, this widespread change is fully taken into account. In *COBAm*, the pronunciation key inside the back cover particularly suggests that this set of words is all transcribed as /w/; there are two keys designated to the symbol /w/, one for *wheat* and *why* and the other for *win* and *wool*. If /h/ is retained, the two keys would be different, i.e. /hw/ (wheat, why and /w/ (win, wool). In fact, the following words are all transcribed /w/, not /hw/, in the entries: whale, what, whatever, wheat, wheel, wheelbarrow, when, where, whether, which, while, whine, whip, whisper, whistle, white, why. In *LAAD2* also, /h/ is absent in all of these items. In *COB5*, on the other hand, /h/ is particularly reserved for American pronunciation, with the notation /hw/: “This shows that some people say /w/, and others, including many American speakers, say /hw/ . . . .” (the Guide, p. xxiv). Thus in *COBAm* (and *LAAD2*), the American pronunciation appears in the way it is expected from the literature.
3.2.7. Vowels before /r/

The vowel inventory before /r/ is worth investigating, since certain contrasts are lost. For example, tense vowels /i/ and /u/ are neutralized with lax vowels /i/ and /o/, respectively, at least in GenAm (Labov et al. 2006—p. 15 Table 2.6 before tautosyllabic /r/ and p. 14 before intervocalic /r/, Wells 1982: 481–485). The following subsections discuss four vowel oppositions before /r/ that are described in the literature as tending to disappear in GenAm.

3.2.7.1. /a/ vs. /A/ before intervocalic /r/

That /a/ and /A/ do not contrast before /r/ is a feature of GenAm (cf. the contrast is retained before intervocalic /r/ in RP): “In many other other [sic] accents, including most GenAm, the two possibilities [ar and Aar] merged, yielding present-day /ar/ in all the words mentioned and giving rise to new rhymes such as furry–hurry, stir it–turret” (Wells 1982: 201).

Our cursory survey of eight words in COBAm that could potentially have /A/ if the vowels did not merge (courage, current, curry, furrow, hurry, occurrence, turrett, worry) finds /ar/ in all words, except in one, hurry, for which both forms /ar/ and /Aar/ are given. COB5 consistently gives /ar/ and /Aar/ (= /ar/ in COBAm’s notation), the latter accompanied by a tag to indicate that it is the American pronunciation. LAAD2 provides both /a/ and /Aar/ in all of the above items except in current, which is assigned only to /a/. This is somewhat surprising, given that it is a dictionary of American English.

3.2.7.2. /e/ vs. /e/, /e/ vs. /e/ before intervocalic /r/

Wells (1982: 480) describes frequent neutralization between /e/ and /e/ before /r/ as being “… recorded by PEAS in western New England and upstate New York and obviously very frequent in the middle and far west”. This results in homophony in pairs of words such as marry–merry, Harold–herald. He further notes GenAm as “often characterized by three-way homophony in sets such as marry–merry–Mary, all /meri/ …” (p. 482). Labov et al. (2006) confirms this three-way homophony by examining two minimal pairs, namely, /et/ Mary vs. /e/ marry, and /e/ marry vs. /ae/ marry. It is unfortunate that the query was not pursued in most areas of the West and Midwest (p. 56, f. n. 6), since these areas are normally included in the GenAm areas geographically. In the areas that the survey covers, they find neutralization (or merger in their term) for both pairs in the great majority of the speakers (p. 54), while “[a] contrast of all three is maintained in the Mid-Atlantic states. Merry and marry are kept apart by a fair number of speakers in New England and the South as well as in Montreal, Quebec” (p. 56, in Map 8.4).

Turning to COBAm, COB5 and LAAD2, we discuss cases before intervocalic /r/. The contrast between /e/ (e.g. marry) and /ae/ (e.g. marry) is maintained in all three dictionaries; they do not show the neutralization observed in the literature. The words we checked are given in (4), showing also the vowels assigned to them.

(4) /e/ and /ae/ are not neutralized before /r/ in COBAm, COB5 and LAAD2

<table>
<thead>
<tr>
<th>COBAm</th>
<th>COB5</th>
<th>LAAD2</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ar/ (COBAm, LAAD2), er (COB5): berry, error, ferry, herald, herring, merit, merry, very</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ar/ (COBS): apparent, carrel, carrot, carry, charity, harridan, harrow, marry, narrow, parity</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

With respect to the opposition between /et/ and /e/, we checked the words represented by square in Wells (1982: 157) under (c), where /r/ is followed by a vowel: aerial, aquarium, area, canary, caring, dairy, fairy, hairy, Mary, parent, Pharaoh, prairie, rarity, various, vary. LAAD2 and COB5 show the neutralization, transcribed as /e/ in LAAD2 and as /ear/ (= /er/ in GenAm) in COB5. In COBAm, all the above words except one show neutralization, transcribed as /ear/, assumed to be interpreted as /er/ in GenAm (see 3.2.1 for interpretation of /a/ before /r/ in COBAm). The one exceptional word is Pharaoh, which has two alternative pronunciations, /e/ and /ear/, in addition to /ear/; /ear/ indicates a potential occurrence of /et/ before intervocalic /r/. However, this might not be a counterexample to neutralization between /et/ and /e/: if the word is syllabified so that /et/ and /e/ do not occur in the same domain, vowel neutralization would not apply, as in compounds such as day room /dei rum/. If this is the case, Pharaoh is not a counterexample to the
neutralization. Given the robust pattern found in other words, we conclude that /ei/ and /e/ are consistently neutralized before intervocalic /r/ in COBAm.

In sum, in all three dictionaries, the possible “three-way homophony” as discussed in the literature is not fully expressed, with merry and Mary being homophones /ˈmɛrɪ/, opposing to marry /ˈmərɪ/. (/er/ and /ər/ may not always sound distinguished in the audio data in COBAm, as discussed later in 3.3.)

3.2.7.3. Two os before tautosyllabic /r/

The literature (e.g. Wells 1982, Labov et al. 2006) describes a widespread merger of non-high back rounded vowels before tautosyllabic /r/ in GenAm, resulting in homophones in pairs such as hoarse—horse and mourning—morning. Specifically, Labov et al. (2006), investigating these two minimal pairs (hoarse—horse, mourning—morning), finds that the distinction has widely disappeared in a large area: “In the middle of the twentieth century, the distinction between /ohr/ and /ohr /[sic]/ in four vs. for, hoarse vs. horse, etc. was maintained by most American speakers, with the exception of the Midland area, centered around Philadelphia, the Mid-Atlantic area, New York, and the Hudson Valley. In the ANAE data, this distinction is made only by a scattering of speakers in Eastern New England, southern Illinois and Indiana, and the Gulf States” (p. 52, in Map 8.2).

In COBAm, the pronunciation key inside the back cover suggests the merger, giving only /xr/. We further researched, in the entries, the two sets of words potentially distinguished, represented by force and north in Wells (1982). We considered words with word-final and preconsonantal /r/s. We only took the strong form when there was a weak form as well (e.g. for, or). The merger was found; all words are shown as /xr/ except for two, for and force, which have /xər/, with the length mark. This length mark appears to be a typo, since there is no use of this diacritic in COBAm in other places. Therefore, we conclude that there is no distinction between hoarse and horse, mourning and morning, etc. in COBAm.

In COBS and LAAD2, a subset was sampled, i.e. words representing the lexical sets (i.e. force and north) and the first two words in each line as listed in Wells (1982: 160, 162). In both dictionaries, these two groups of words are assigned to the same vowel (/xr/ in COBS, /x/ in LAAD2).

In sum, none of the three dictionaries express the distinction, which conforms to what the literature finds.

3.2.8. Lexical incidences

There may be differences of pronunciation which are lexical and not systemic. Wells (1982: 126–127) lists examples that differ between GenAm and RP, though he also notes that the “GenAm”/“RP” categorization is not cut and dry: GenAm forms may be heard in England and vice versa; he often qualifies the GenAm forms as “usually” and “often”. That said, we take the “GenAm” pronunciation as the reference point to examine whether it is found in COBAm, as compared with COBS and LAAD2.

The words are address, advertisement, anti-, ate, ballet, Bernard, beta, borough, clerk, cuckoo, depot, detail, docile, erase, figure, from (strong from), herb, inquiry, iodine, laboratory, leisure, lever, lieutenant, massage, neither, nonsense, of (strong form), omega, primarily, process (n.), progress, quinine, record (n.), schedule, shone, suggest, thorough, tomato, vase, vermouth, was (strong form), wrath, Z.

About half of these words have only the GenAm form in all three dictionaries (in COBS, forms marked as AM are considered). For the rest, there are some notes to be made. First, in the two American English dictionaries, COBAm and LAAD2, as one may expect, the GenAm form appears for all items, while in COBS, there are three items (and possibly anti- as well; see Note 17) that are not given the GenAm form but instead have the RP form only. Overall, the two American English dictionaries indicate more American pronunciation compared to COB5.

Second, there are items for which both RP and GenAm forms are given. This is expected from Wells’s note on the non-categorical nature of the labels “GenAm” and “RP” as discussed above, though in COBAm, the variation may mot be within GenAm but between GenAm and RP, as discussed before. COBAm has slightly fewer items of this kind than
Third, if advertisement and anti- are excluded (see Note 17), COBAm only provides the forms included in Wells (including cases also accompanied by an RP form), while LAAD2 has at least three instances that represent the third possibility, i.e. neither Wells's GenAm nor RP: iodine has /aɪdɪn/ along with GenAm /aɪdɪn/ (cf. RP /aɪdɪn/); progress has /ˈprəgres/ along with GenAm /ˈprəgres/ (cf. RP /ˈprəgres/); vase has /veɪs/ along with GenAm /vɛs/ and RP /væs/. In Merriam-Webster's Collegiate Dictionary (2003), another American English dictionary, iodine does not have this third form, while progress and vase do. /rəs/ is listed first in progress. /veɪs/ for vase is noted as a form found in “Canada usu[ally] & US also”. Thus, at least for progress and vase, LAAD2 describes lexical variation more accurately than COBAm.

3.3. Pronunciation on the COBAm CD-ROM

In this section, we examine whether the pronunciation on the CD-ROM of COBAm is consistent with the transcription given in the paper version of the dictionary and also in turn with the literature. (See discussion in Section 7 for functional aspects of the CD-ROM, whose facets include those related to pronunciation as well as others.) In order to do this, we select a subset of the diagnostic features discussed in 3.2, that is, merger of low back vowels, ask words, Yod dropping, reduction of the /hw/ onsets, and neutralization between /e/ and /æ/ before intervocalic /r/.

We consider the same words as we did above in reviewing the paper version of the dictionary, except in merger of low back vowels, ask words, Yod dropping, reduction of the /hw/ onsets, and neutralization between /e/ and /æ/ before intervocalic /r/.

Regarding the contrast between /æ/ and /a/, recall that these vowels are generally assumed to have merged to /a/ in GenAm in the literature. The vowels in a near minimal pair pod—pa both sound unrounded, though the vowel in pa sounds slightly lower than that in pod. Note that from the transcription allotted to pod, /n/, the unrounded quality is not expected. With respect to the contrast between /æ/ and /a/, there are at least a couple of minimal pairs that are pronounced with very little difference on the CD-ROM, knotty—naughty and don—dawn. The vowel could be transcribed [g], that is, slightly higher than, and not as back as, cardinal [a]. The lips do not sound as if they are actively involved in the articulation, i.e. neither actively rounded as cardinal [a] nor actively unrounded as cardinal [a].

The closeness in the vowel quality in these (near) minimal pairs should be compared with the same words in COB5; COB5 seems to provide only RP pronunciation for these items, where the three vowels sound quite described in the paper version, including the variation in avenue and residue, except in one word. For during, [i] is heard after [d], whereas it is dropped in the transcription.

A little more deviation In ask words, the deviation from the paper version is slightly greater. It is still the case that all but one (Slav) have at least the GenAm form with /æ/. However, the CD-ROM records many more instances of the alternative /a/ in addition to /æ/ than the paper version does. Words added to the list in (3b) above are: cask, casket, cast, castle, castor, chaff, chance, chancel, chancellor, chant, clasp, class, command, commando, craft, fasten.

Interesting deviations Finally, there are two features that are not indicated in the paper version but are discussed in the literature and appear to be recorded on the CD-ROM. These are the loss of certain vowel contrasts, i.e. low back vowels /o/ (lot etc.) vs. /a/ (palm etc.) vs. /æ/ (cloth, thought, etc.), and /e/ (hairy, merry, etc.) vs. /æ/ (harry, marry, etc.) before intervocalic /r/. On the CD-ROM, the vowels in these sets of words do not always sound different. We discuss the possibility of the low back merger first, and the possible neutralization between /æ/ and /æ/ before intervocalic /r/ next.

Regarding the contrast between /æ/ and /a/, recall that these vowels are generally assumed to have merged to /a/ in GenAm in the literature. The vowels in a near minimal pair pod—pa both sound unrounded, though the vowel in pa sounds slightly lower than that in pod. Note that from the transcription allotted to pod, /n/, the unrounded quality is not expected.
distinct from each other. With respect to the rounding dimension only, /ɒ/ (don, knotty, pod) is rounded, /ɑː/ (dawn, naughty) is more rounded, and /a:/ (pa) is unrounded. Compared to this, the vowels in a pair are fairly similar to each other in COBAm. If this little difference in the two pairs is not significant, it suggests that the three-way contrast employed in the paper COBAm might not be present in the production data on the CD-ROM, which is not surprising from what the literature describes for GenAm, i.e. loss of contrasts in these pairs.

An acoustic analysis was conducted for the pairs knotty /ɒ/—naughty /ɑː/ and don /ɒ/—dawn /ɑː/. Assuming these words are spoken by the same speaker, the formants are comparable without normalization. We used Adobe Audition 1.5 to transfer the MPEG files to WAV files so that they could be analyzed by Praat (version 4.6.01, Paul Boersma and David Weenink). The mean formant values were measured for a period of 100 ms in the vowel. In the knotty—naughty pair, the 100ms duration was taken as 50 ms forward and 50 ms backward from the approximate midpoint during which the formants appear steady. In the don—dawn pair, since F1 is not detected after a certain point in the vowel, 100ms was taken until the point at which F1 detection starts failing. The results are plotted in Figure 3.1.

![Figure 3.1 F1 and F2 in the stressed vowels in knotty-naughty and don-dawn](image)

Figure 3.1 F1 and F2 in the stressed vowels in knotty-naughty and don-dawn


In both pairs, F2 is almost the same, which suggests that the backness of the tongue is almost the same in a given pair. F1 differs slightly, with /ɑː/ being slightly lower than /ɒ/ (by 50 Hz in the knotty-naughty pair and by 37 Hz in the don-dawn pair). The direction is consistent: F1 of /ɑː/ is always lower than F1 of /ɒ/. This is expected from what the transcriptions suggest: /ɑː/ is higher than /ɒ/ in vowel height. Since F1 correlates inversely with vowel height (i.e. the higher the F1 the lower the tongue height), F1 in /ɑː/ is expected to be lower than F1 in /ɒ/. It would be interesting to test whether this difference is also consistent in other minimal pairs, and, if so, whether this small difference is statistically significant and perceptually significant, i.e. whether native speakers perceive the approximately 45 Hz of F1 difference in these pairs.

Regarding the opposition between /e/ and /æ/ before intervocalic /r/, while it is transcribed to be present in the paper COBAm, there are (near) minimal pairs on the CD-ROM in which impressionistic analysis again finds neutralization. These pairs include hairy /e/—harry /æ/ and caring /e/—carry /æ/. It is possible to compare /æ/ in this neutralization context, i.e. before /r/, with /e/ in non-neutralizing context, since the contrast exists elsewhere. For example, /æ/ in hairy can be compared with /e/ in had, likewise /æ/ in carry with /e/ in cad. This comparison finds quite a difference, with the /æ/ in had and cad realized around /e/ or lower, while /e/ before /r/ sounds as high as /e/ before /r/, hairy [e]—harry [æ] and caring [e]—carry [æ]. To these pairs, the pair merry /e/—marry /æ/ may be added: the vowel in marry is not as front as the vowel in merry, but it is as high, merry [e]—marry [æ], which is to be compared to the much lower [æ] in mad.

Acoustic analysis of these sets of words found possible neutralization between /e/ and /æ/ before /r/. The same method as described above was used to convert the sound files on the CD-ROM to WAV files. For measuring F1 and F2, in the vowels in cad, had and mad, the mean values for 100ms were taken around the portion where F1 and F2 appear steady with a regular wave form. For the vowels before /r/, since F2 (and F3) decreases throughout toward /r/ (which is expected in a vowel surrounding /r/), the mean F1 and F2 were taken from the point of F2
maximal to the point of F2 minimal. The results are shown in Figure 3.2, where F1 and F2 are plotted.

![Figure 3.2 F1 and F2 in stressed vowels in hairy, harry, had, caring, carry, cad, merry, marry, mad](image)

Figure 3.2: F1 and F2 in stressed vowels in hairy, harry, had, caring, carry, cad, merry, marry, mad

There are two recognizable groups in Figure 3.2: one with /ɛ/ and /æ/ before /r/ (hairy, harry, caring, carry, cad, merry, marry, mad) and the other with /æ/ before /d/ (had, cad and mad). In terms of the height (acoustic correlates in F1), in each set (hairy–harry–had, caring–carry–cad, and merry–marry–mad), /æ/ is raised as high as /ɛ/ before /r/, since the F1 value of /æ/ before /r/ is as low as F1 of /ɛ/ before /r/, compared to F1 of /æ/ before /d/, which is much higher. In terms of backness (acoustic correlate in F2), /æ/ is as back as /ɛ/ before /r/, while /æ/ before /d/ is much more front. There is a larger difference between merry and marry in F2 (by 83 Hz) than the difference in the other pairs (38 Hz in hairy–harry, 54 Hz in caring–carry), which is also detected in the impressionistic analysis as discussed above. In order to see if this shows incomplete neutralization, more tokens must be examined. Nonetheless, it is observable that the F2 value of /æ/ in marry is much closer to /ɛ/ in merry than to /æ/ in mad.

Thus, as far as these three pairs are concerned, the production data on the CD-ROM suggest neutralization between /ɛ/ and /æ/ before intervocalic /r/ both impressionistically and acoustically; /æ/ is as high as /ɛ/ in this context. This is expected from the literature, but is not incorporated in the transcriptions in the paper version of COBAm. This might be confusing to the user of the dictionary, since those words that are assigned to different vowel symbols in the paper version sound nearly the same on the CD-ROM.

3.4. Some notational discrepancies in COBAm

There are several typos in the paper version of COBAm. Most of them would not be confusing to many users of the dictionary, because the wrong symbol is not used in other places (e.g. tap “c” for /r/ in the Guide) or because the difference is of a kind which only linguists would be concerned about (e.g. script “a” in diphthongs /ai/ /au/ in the list of vowels inside of the back cover). We would like to draw attention to one of these, however, since it might be of greater importance than the others to dictionary users. In the note about the symbol /ou/ in the Guide, /u/ is incorrectly given as the RP counterpart of GenAm /ou/. It is generally thought to be /au/ (as in COB5’s Guide), not /u/.

(Section 3 by Hirayama)

4. Definitions

In this section the treatment of meaning in COBAm, semantic and pragmatic, will be discussed mainly through the comparison of definitions and labels in COBAm with those in COB5. This section is divided into the following subsections: defining vocabulary (4.1), sense description (4.2), and labels and pragmatics (4.3).

For the comparison between the two dictionaries, all the entries on the following thirty pages of COBAm are compared with the corresponding entries from COB5: pp. 100–1, 200–1, 300–1, 400–1, 500–1, 600–1, 700–1, 800–1, 900–1, 1,000–1, 1,100–1, 1,200–1, 1,300–1, 1,400–1, 1,500–1. When a headword is not included in either of the dictionaries, its definition does not count. Other entries are also examined whenever necessary.
The results of our examination show that there are no remarkable differences between *COB5* and *COBAm* except several regular replacements of words used in definitions and the incorporations of labels. Details will be given in relevant subsections which follow.

### 4.1. Defining vocabulary

So far as definitions are concerned, the inclusion of the list of defining vocabulary (henceforth DV) in *COBAm* is among the most striking differences between *COBAm* and *COB5*. However, this does not really seem to reflect some change in their policies on the uses of DV, because the majority of definitions from *COBAm* are almost identical to the corresponding definitions from *COB5*. It is also to be mentioned that despite *COBAm*'s claims to have used DVs, words outside of the DVs are used in definitions without being marked in any way. We will discuss problems with *COBAm*'s explanation of its policy on the use of DV in 4.1.1 and the results of our examination of its use of DV in 4.1.2.

#### 4.1.1. *COBAm*'s policy on its use of DV

It is now customary that EFL dictionaries claim to achieve easier definitions written with DV. At the same time, it is often pointed out that lexicographers’ ways of using DV are not always clear to the user (Fox 1989, Minamide 1995), and *COBAm* is no exception. The guide to definitions in *COBAm* runs as follows:

... the definitions are written in full sentences, using vocabulary and grammatical structures that occur naturally with the word being explained. This enables us to give a lot of information about the way a word or meaning is used by speakers of the language. Whenever possible, words are explained using simpler and more common words. This gives us a natural defining vocabulary with most words in our definitions being among the 2,500 most common words of English (p. xiii)

According to this explanation, definitions in *COBAm* are written with words which “occur naturally with” definienda, which may be why they call their DV “natural”. More importantly the lexicographers explicitly state that they tried to define headwords using “simpler and more common words” but only “whenever possible”. They also say that “most words” in their definitions are “among the 2,500 most common words of English”, which implies that the lexicographers are in principle allowed to use any word even beyond the scope of the DV. If *COBAm* uses any non-DV word in definitions without marking it, the dictionary cannot claim to use DV in its strict sense.

It has often been pointed out that the actual size of DV is usually larger than it looks because all the word forms and phrases which are used as part of DV are not necessarily listed (Herbst 1986, Kirkpatrick 1985). We may well associate this with commercialism; the smaller their DVs appear, the easier and better their products look, which will greatly affect the sales of their products. In fact, while the COBUILD lexicographers merely mention the fact that the majority of words used in their definitions belong to the 2,500 most frequent words in English, 3,221 words are given on the DV list (pp. 1575-84).1)

Moreover, *COBAm* lists no phrasal verbs or idiomatic expressions on the DV list except a few compound nouns, but it is questionable whether it is possible not to use them when writing dictionary definitions. Some scholars are against the use of phrasal verbs and/or idioms consisting of DV items as part of DV because even though their components are included in DV, their total meanings as phrases are not transparent enough (Jansen et al. 1987).

Too strict an application of DV will bring about undesirable effects such as unnatural definitions, so there is nothing wrong with their use of non-DV items sometimes. However, other EFL dictionaries usually mark their exceptional uses in small capital letters when they use words from outside of their DVs (*OALD7* (p. R99), *LAAD2* (p. 1852)). As the above guide does not give us any more information on their policy toward the use of non-DV items, in the next subsection we will return to this point with the results of our examination of how the lexicographers use the DV in *COBAm*.

It is also problematic that *COBAm* does not sufficiently tell what their criteria are for selecting the DV items on the list. There is a possibility that...
COBAm's DV is the same in nature as that of Collins COBUILD Essential English Dictionary (1988) whose DV is the collection of words actually used in its definitions more than ten times (Fox 1989). If this is the case, COBAm's DV does not affect the wording of its definitions because it was not selected prior to the lexicographers’ writing definitions but collected after the definitions had been written.

DV has become an important selling point of EFL dictionaries today as it is believed to help foreign learners of English to understand definitions more easily and correctly. As long as they claim to achieve easier definitions at all with the use of DV, they should therefore let the user know how they use it in their definitions. In the next subsection we will investigate COBAm's use of DV.

4.1.2. The use of DV

As mentioned in the previous subsection, COBAm does not tell us how it actually uses the DV. Looking at definitions from COBAm, however, we find that the dictionary uses words not listed without marking them. Below are examples:

A **crescent** is a curved shape that is wider in the middle than at its ends, like the shape of the moon during its first and last quarters. It is the most important symbol of the Islamic faith. *(crescent)*

**Manure** is animal feces, sometimes mixed with chemicals, that is spread on the ground in order to make plants grow healthy and strong. *(manure)*

Two words not listed as part of the DV, “curved” for **crescent** and “feces” for **manure**, are used in the above definitions without marking them. *(crescent)*

Although COBAm does not list any phrasal verb in its DV list, in the definition of **abbreviation**, a phrasal verb, “leave out” is used. In that of **happen**, two idiomatic expressions are used, neither of which is listed: “be the case” and “by chance”.

We can safely say that COBAm does not use DV in its strict sense. If they claim to adopt DV to help the user understand definitions more easily and correctly, what they need to do is simply give the user the wrong impression that all the definitions in COBAm are written with the listed DV words alone. Of the two non-DV items in the above definitions, “curved” and “feces”, the latter may be less familiar to ordinary EFL learners because of its infrequent and technical nature. Although COBAm states that the lexicographers tried to use the simplest words possible, their claim may not always be true. We are not saying that they should stick to the use of easier words than those being defined, which will be almost impossible. Especially when defining basic vocabulary items lexicographers cannot but use less frequent and therefore more difficult words. Again, we would argue that if they claim to achieve easier definitions with the DV, they must let the user know how and what DV items they use in their definitions.

The same criticism of COBAm's use of non-DV items also applies to its use of idiomatic expressions:

An **abbreviation** is a short form of a word or phrase, made by leaving out some of the letters or by using only the first letter of each word. *(abbreviation)*

If you **happen** to do something, you do it by chance. If it **happens that** something is the case, it occurs by chance. *(happen)*

Although COBAm does not list any phrasal verb in its DV list, in the definition of **abbreviation**, a phrasal verb, “leave out” is used. In that of **happen**, two idiomatic expressions are used, neither of which is listed: “be the case” and “by chance”.

We can safely say that COBAm does not use DV in its strict sense. If they claim to adopt DV to help the user understand definitions more easily and correctly, what they need to do is simply give the user the wrong impression that all the definitions in COBAm are written with the listed DV words alone. Of the two non-DV items in the above definitions, “curved” and “feces”, the latter may be less familiar to ordinary EFL learners because of its infrequent and technical nature. Although COBAm states that the lexicographers tried to use the simplest words possible, their claim may not always be true. We are not saying that they should stick to the use of easier words than those being defined, which will be almost impossible. Especially when defining basic vocabulary items lexicographers cannot but use less frequent and therefore more difficult words. Again, we would argue that if they claim to achieve easier definitions with the DV, they must let the user know how and what DV items they use in their definitions.

The same criticism of COBAm's use of non-DV items also applies to its use of idiomatic expressions:

An **abbreviation** is a short form of a word or phrase, made by leaving out some of the letters or by using only the first letter of each word. *(abbreviation)*

If you **happen** to do something, you do it by chance. If it **happens that** something is the case, it occurs by chance. *(happen)*

Although COBAm does not list any phrasal verb in its DV list, in the definition of **abbreviation**, a phrasal verb, “leave out” is used. In that of **happen**, two idiomatic expressions are used, neither of which is listed: “be the case” and “by chance”.

We can safely say that COBAm does not use DV in its strict sense. If they claim to adopt DV to help the user understand definitions more easily and correctly, what they need to do is simply give the user the wrong impression that all the definitions in COBAm are written with the listed DV words alone. Of the two non-DV items in the above definitions, “curved” and “feces”, the latter may be less familiar to ordinary EFL learners because of its infrequent and technical nature. Although COBAm states that the lexicographers tried to use the simplest words possible, their claim may not always be true. We are not saying that they should stick to the use of easier words than those being defined, which will be almost impossible. Especially when defining basic vocabulary items lexicographers cannot but use less frequent and therefore more difficult words. Again, we would argue that if they claim to achieve easier definitions with the DV, they must let the user know how and what DV items they use in their definitions.

The same criticism of COBAm's use of non-DV items also applies to its use of idiomatic expressions:

An **abbreviation** is a short form of a word or phrase, made by leaving out some of the letters or by using only the first letter of each word. *(abbreviation)*

If you **happen** to do something, you do it by chance. If it **happens that** something is the case, it occurs by chance. *(happen)*
An analysis of Collins COBUILD Advanced Dictionary of American English

4.2. Sense description

There are no remarkable differences between sense descriptions, or definitions in COB5 and COBAm. However, as COBAm is designed for learners of American English, Briticisms used in definitions in COB5 are regularly replaced with the corresponding Americanisms in COBAm. We will report on these replacements in 4.2.1 and other changes made in 4.2.2. The arrangement of senses will be discussed in 4.2.3.

4.2.1. Replacements of Briticisms with Americanisms

COBAm adopts the full sentence definition style as other COBUILD dictionaries do, and the vast majority of definitions in COB5 and COBAm are identical except that Briticisms used in definitions in COB5 are replaced with the corresponding Americanisms. On the sample pages there are 519 words with 1,053 senses and phrases common to COB5 and COBAm. 1,022 out of the 1,053 senses and phrases are defined in exactly the same way except for the replacements. As COBAm is designed for learners of American English, these changes are reasonable. The replacements are carried out on the levels of spelling and lexical items.

British spellings used in definitions of COB5 are systematically replaced by the corresponding American spellings in COBAm. An example is given below:

COB5: If you use bask in someone's approval, favour, or admiration, you greatly enjoy their positive reaction towards you. (bask [2]; italics are ours.)

COBAm: If you use bask in someone's approval, favor, or admiration, you greatly enjoy their positive reaction toward you. (bask [2]; italics are ours.)

In the above definitions British spelling, "favour", and "towards" (which is not strictly British spelling but preferred in British English) are replaced with their American counterparts in COBAm, "favor" and "toward" respectively. These replacements of spelling are regularly made throughout definitions in the two dictionaries. Other examples include "programme" changed to "program" (e.g. introduce 4) and "behaviour" to "behavior" (e.g. bask).
Words which are typically used in British English are also systematically replaced with their American equivalents in COBAm. Below is an example:

**COB5** A dumb waiter is a lift used to carry food and dishes from one floor of a building to another. (dumb waiter; italics are ours.)

**COBAm** A dumbwaiter is an elevator used to carry food and dishes from one floor of a building to another. (dumbwaiter (dumb waiter); italics are ours.)

While COB5 uses “lift” for the definiendum, it is replaced with its American equivalent, “elevator” in COBAm. Similar examples are legion: “film” replaced by “movie” (remaster) and “flat” by “apartment” (studio 4). It is to be mentioned that the replacements are carried out according to a sense of a word used:

**COB5** You can use celluloid to refer to films and the cinema. (celluloid; italics are ours.)

**COBAm** You can use celluloid to refer to movies. Celluloid is a type of plastic formerly used for making photographic film. (Italics are ours.)

In the above definitions, “films and the cinema” in COB5 is replaced with its American equivalent, “movies” in COBAm. Importantly, COBAm adds encyclopedic information on the headword in the second sentence in its definition, where “film” is used referring to a kind of plastic. This shows that “film” is only replaced with “movie” when it is used meaning movie.

Similar changes are made when a particular word is almost exclusively used in either of the varieties of English or when it is preferred. Below is an example:

**COBAm** A mantelpiece is a wood or stone shelf which is the top part of a border around a fireplace. (mantelpiece; italics are ours.)

While “around” is used in the above definition, “round”, which may be preferred in British English, is used in COB5. However, these replacements are not consistently made. For instance, both COB5 and COBAm use “around” in their definitions of marauder, marauding, and shawl. Similarly, “autumn” is replaced with “fall” in the definition of maple, but not in shed (Sense 3).

Replacements of Briticisms are also carried out in a more subtle way. According to Longman Grammar of Spoken and Written English (1999: 616), relative clauses with “that” are preferred to those with “which” in American English. On the sample pages there are several definitions where the relative pronoun “which” used in COB5 is replaced by “that” in COBAm:

**COB5** You use invasive to describe something undesirable which spreads very quickly and which is very difficult to stop from spreading. (invasive A; italics are ours.)

**COBAm** You use invasive to describe something undesirable that spreads very quickly and that is very difficult to stop from spreading. (invasive 1; italics are ours.)

Other examples include the definitions of pocket (Sense 1) and pointer (Sense 1). Although this replacement of “which” is not always made, we may safely say that in most cases there are systematic replacements of Briticisms by Americanisms throughout definitions in the two dictionaries.

### 4.2.2. Other changes

While the vast majority of definitions in COB5 and COBAm are exactly the same except for the replacements of Briticisms, there are several other changes made. These changes can be roughly classified into the following three types: attempts to achieve more concise and/or readable definitions, additions of more information, and reflections of political correctness.

Although our judgements on readability depends on various factors even including a personal preference, the following definition from COBAm can be slightly more readable, at least more concise, compared to the corresponding one from COB5:

**COB5** Fixtures are pieces of furniture or equipment, for example baths and sinks, which are fixed inside a house or other building and
which stay there if you move. (fixture 1)

COBAm  **Fixtures** are fittings or furniture which belong to a building and are legally part of it, for example, a bathtub or a toilet. (fixture)

The second relative clause in COBS's definition, "which stay there if you move", seems rather complicated to a foreign learner as it includes another clause where a concessive use of "if" is used. On the other hand, COBAm uses a simple phrase instead, "legally part of it", and inserts a comma between the phrase, "for example" and the following two examples, which may help the user parse the structure of the definition more easily. This insertion of a comma is also found in the definitions of such words as **pocket** (Sense 9 in COBAm), **sharp** (Sense 5) and **shed** (Sense 2). However small this insertion looks, this can be regarded as COBAm's attempt at achieving more readable definitions.

There are also cases where COBAm provides more information than COBS does:

**COBS**  If you **hang up** or you **hang up** the phone, you end a phone call. If you **hang up** someone you are speaking to on the phone, you end the phone call suddenly and unexpectedly. (hang up 2)

**COBAm**  If you **hang up** or you **hang up** the phone, you end a phone call. If you **hang up** someone you are speaking to on the phone, you end the phone call suddenly and unexpectedly, usually because you are angry or upset with the person you are speaking to. (hang up 2)

COBAm adds a clause explaining the typical reason for one's hanging up the phone so that the user could grasp the meaning of the phrase more correctly. The definition of **celluloid** discussed in 4.2.1 is another example where COBAm gives more information on the headword, encyclopedic information in that case. We can find more examples in the definitions of **introspective** and **tsunami**.

Similarly, COBAm adds a piece of information by using a phrase, "In economics" to Sense 2 of **manufacture**: "In economics, manufactures are goods or products which have been made in a factory". As COBS does not specify the field where this particular sense is used, this addition may be a step forward. However, this could be dealt with as effectively or slightly better if a relevant label like [ECONOMICS] were used. COBAm does not use the label perhaps because it does not have the label. If more labels had been introduced and used properly to specify a field where a particular sense of a headword is used, it would help the user locate the sense they are looking for more easily with the labels as a signpost (see 4.3.2. for more discussions about COBAm's use of labels).

The last point to be made concerning changes of wording of definitions is political correctness. For instance, while COBS defines **truncheon** as follows: "A truncheon is a short, thick stick that is carried as a weapon by a policeman", "a policeman" is replaced with "a police officer" in COBAm.

Below is another example where "man" is replaced by "person":

**COBS**  A **truss** is a special belt with a pad that a **man** wears when **he** has a hernia in order to prevent it from getting worse. (truss 2; italics are ours.)

**COBAm**  A **truss** is a special belt with a pad that **person** wears when **they** have a hernia in order to prevent it from getting worse. (truss II; italics are ours.)

Since the use of "man" is not preferred nowadays when referring to a person of unknown sex, "a man" in COBS is replaced by "a person" in COBAm. Moreover, "he" is also replaced with "they" even violating a rule of traditional grammar; i.e. the agreement of number. These replacements reflect the true picture of the English language as used today, and thus they may also be among welcome features of COBAm. Yet, since the above use of "they" does not agree with its antecedent in number, some foreign learners may find it difficult to understand the definition if they do not know this particular use of the pronoun. 2)

In addition to the replacements of Briticisms in definitions, there are some changes made between definitions in COBAm and COBS. On the sample pages, there is also a case where a grammatical mistake or typo in a definition in COBS is corrected in COBAm, e.g. **sharp-tongued**. While the vast majority of definitions from COBAm and COBS are the same, COBAm seems to have attempted to achieve better sense descriptions.
4.2.3. Arrangement of senses

Neither COBS nor COBAm states its policy on how different senses of a headword are arranged in an entry. As far as the sample pages are concerned, there are no fundamental differences between their sense arrangements. Although there are several promotions and demotions of senses in entries such as those for credit and shave, no notable regularity can be found. Thus we compared all the numbered senses of take: ① take and ② take from COBS and COBAm, excluding those for phrasal verbs.

COBS has 46 senses for the headword, and COBAm has 44; while one sense (Sense 43) has been added in COBAm, four senses (Senses 27, 32, 40 and 41) have been deleted from COBS. As a result, other senses in COBS have been automatically promoted except that Senses 17 and 18 in COBS have been respectively demoted to Senses 35 and 34 in COBAm.

With respect to the changes made in the sense arrangements in COBS and COBAm, it would be reasonable to assume that those senses which are used in British English were deleted or demoted and that American senses were entered or promoted, because, as we saw in 4.2.1, COBAm has a strong inclination toward American English. However, among the four deleted senses, only one sense (Sense 32) is marked with a label, [mainly BRIT]. Concerning the two demoted senses, one (Sense 17 in COBS) is marked with [mainly BRIT], but the other with [mainly AM]. We cannot explain the promotions and the demotions from the fact that COBAm focuses on American English.

Another possibility is that the changes were made in the process of updating the descriptions of the headword by using the latest findings from large corpora such as frequency data. Yet, so long as we compared the treatments of the two demoted senses in COBAm, CALD2, LDOCE4 and OALD7, the two senses may not deserve the demotions. COBAm might have arranged senses according to different criteria from those that other dictionaries use. Or, this is simply because the dictionary is based on data from the Bank of English which other dictionaries do not use. We will return to this point in the next subsection.

4.2.4. Coverage of senses

There are several additions and deletions of senses of headwords common to COBS and COBAm. On the sample pages, only ten senses have been added in COBAm, whereas 104 senses have been deleted from COBS. It will be clear that COBS has a wider coverage as far as senses of headwords are concerned.

Concerning the promotions and demotions there is one tendency; senses arranged at or near the end of an entry in COBS were deleted in COBAm. As we have seen in the previous subsection, these deletions may not be only due to the lexicographers’ attempts to Americanize COBS. Only eight out of 104 senses deleted from COBS are marked as [BRIT] or [mainly BRIT]. There are also several senses which are marked as [AM] or [mainly AM] in COBS but not contained in COBAm: Sense 2 of credit transfer and Sense 5 of stump. Importantly, forty out of the above 104 deleted senses are listed at the end of the entries in COBS. Sometimes several senses arranged in succession have been deleted together when they are listed at the end of their entries. If we count them, 61 out of the 104 senses have been deleted from COBS in this way.

For example, under the entry for remember in COBS, eight senses are listed and the last three Senses, 6, 7 and 8 have been deleted in COBAm. LAAD2 has nine senses for the verb, and Senses 6 and 9 in the dictionary correspond to Senses 6 and 7 in COBS respectively. Sense 8 in COBS is not contained in LAAD2, and this sense will not be discussed here. Because both dictionaries list the two senses near the end of the entry for remember, they must regard the senses less important in some way, say in terms of frequency. However, it is not to be overlooked that while COBAm does not have them, LAAD2, another EFL dictionary for advanced learners of American English published in the same year as COBAm, still has the two senses. Other EFL dictionaries, CALD2, LDOCE4 and OALD7, also have the two senses. As we mentioned in the previous subsection, COBAm’s criteria for selecting senses to include seems different from those of other EFL dictionaries. Moreover, COBS, which was published just one year prior to COBAm and also uses the Bank of English, has the senses. The deletions cannot be accounted for by
COBAm's use of the Bank of English.

As far as individual senses of headwords common to COB5 and COBAm are concerned, COBAm has a narrower coverage, and COBAm has sometimes automatically deleted those senses listed at or near the bottom of the entries without carefully considering its possible importance. COBAm's criteria for selecting which sense to include are not really clear.

4.3. Labels and pragmatics

In this subsection COBAm's use of labels, geographical, style and pragmatic, will be discussed. While there are no fundamental differences between the uses of the labels in COB5 and COBAm, COBAm's use of the pragmatic labels seems to have been degraded mostly due to the abolishment of the Extra Columns; embedded within each entry, the labels have become less conspicuous.

This subsection is divided into two: the use of the geographical and style labels (4.3.1), and the use of the pragmatic labels together with the treatment of pragmatic information (4.3.2). As the style labels and the pragmatic labels are sometimes closely connected, we will consider them together in 4.3.2.

4.3.1. Geographical and style labels

The lists of geographical labels in COB5 and COBAm are almost the same except that the former lists more labels like IRISH. As for the style labels, COBAm lists fewer labels by two; while RUDE, VERY OFFENSIVE and VERY RUDE have been omitted from the list in COB5, VULGAR was added in COBAm. Comparing the lists of the style labels in the two dictionaries more closely, we will note that VERY RUDE seems to have been replaced with VULGAR in COBAm as their explanations of the labels are identical.

In order to see how the labels are used in COBAm, we have examined all the entries in COB5 and COBAm. The results show that most of the words labeled RUDE in COB5 are regularly labeled VULGAR in COBAm: boob, fart and go to hell. At the same time VERY RUDE is also systematically incorporated into VULGAR (e.g. mother fucker, prick and son of a bitch). Interestingly, as for mother fucker and son of a bitch, OFFENSIVE and a pragmatic label, DISAPPROVAL are added in COBAm. We will return to this point in the next subsection. Concerning VERY OFFENSIVE, wog, which COB5 gives as an example of headwords with the label, is not included in COBAm, but the label is still used in COBAm: coon and nigger. COBAm does not list the label perhaps because there are not many words labeled VERY OFFENSIVE.

Apart from the incorporation of RUDE and VERY RUDE, the uses of labels in COB5 and COBAm are basically the same, and the incorporation may reflect an attempt of COBAm to simplify its use of the labels.

4.3.2. The pragmatic labels

Pragmatics started attracting attention in EFL lexicography when LDOCE2 was published in 1987. However, COB2 is arguably the first EFL dictionary which incorporated pragmatics into EFL lexicography in a systematic way, introducing a label, PRAGMATICS throughout its entries. COB3 more elaborately classified the label into seven, [approval, disapproval, emphasis, feelings, formulae, politeness and vagueness]. Its successors, COB4, COB5 and COBAm, also employ these seven labels.

There are no fundamental differences between COB5 and COBAm with respect to their uses of the pragmatic labels. Nevertheless, it will be fair to say that COBAm's use of the pragmatic labels is greatly degraded due to the abolishment of the Extra Columns. The labels in COBAm are embedded within each entry without the square frame and thus far less conspicuous. Although the lexicographers put them in small capital letters, other labels are also presented in that way. Considering the fact that EFL dictionaries today contain a variety of information, lexicographers must employ a method to help the user recognize the kind of information they are looking for more easily.

It is also to be mentioned that at every revision the COBUILD dictionaries have been reducing the explanation for pragmatics in the front matters. While COB2 spent four pages explaining what pragmatics is and how important it is in the EFL context, COB5 spent only one page; COBAm spends less than one, which even gives us the impression that the COBUILD lexicographers do not pay much attention to pragmatics any
As we saw in the previous subsection, there are several additions of pragmatic labels in COBAm. Below are examples:

If someone calls a person, usually a man, a **motherfucker**, they are insulting him in a very unpleasant way. [mainly AM, OFFENSIVE, VULGAR, DISAPPROVAL] (**motherfucker**)

If someone is very angry with another person, or if they want to insult them, they sometimes call them a **son of a bitch**. [INFORMAL, OFFENSIVE, VULGAR, DISAPPROVAL] (**son of a bitch**)

While COBAm adds DISAPPROVAL to the above definitions, there are no pragmatic labels in COBS’s corresponding definitions. These additions may agree with COBAm’s explanation of pragmatics: “People use words to do many things: give invitations, express their feelings, emphasize what they are saying, and so on. The study and description of the way in which people use language to do these things is called **pragmatics**” (p. xvi). Clearly, COBUILD lexicographers equate pragmatics with speech acts. Thus, it will be reasonable for them to regard the use **motherfucker** or **son of a bitch** as belonging to the domain of pragmatics because to use them can carry out a speech act; that is to say, to insult a person spoken or referred to in this case. Interestingly, there are two labels assigned to them which are similar to each other: OFFENSIVE and DISAPPROVAL. We may naturally wonder what their differences are. COBAm explains them as follows:

**OFFENSIVE**: likely to offend people, or to insult them; words labeled OFFENSIVE should therefore usually be avoided (p. xv)

**disapproval**: used to show that you disapprove of the person or thing you are talking about (p. xvi)

If pragmatics is solely concerned with how people use words to carry out an act as COBAm defines it, OFFENSIVE can also belong to pragmatics rather than to style because to use words or expressions whose use is “likely to offend people, or to insult them” can naturally offend and insult a person. More importantly, to use **motherfucker** and **son of a bitch** would arguably not merely show our disapproval of the person we are talking about.

Below are examples where pragmatic labels, FEELINGS and EMPHASIS are added in COBAm:

If you tell someone to **go to hell**, you are angrily telling them to go away and leave you alone. [INFORMAL, VULGAR, FEELINGS] (**hell** 4)

If you say that someone can **go to hell**, you are emphasizing angrily that you do not care them and that they will not stop you doing what you want. [INFORMAL, VULGAR, EMPHASIS] (**hell** 5)

If we use the former expression with someone, we can tell them to go away and leave us alone. We can also tell them that we are angry. The addition of FEELINGS thus agrees with COBAm’s definition of pragmatics, but to use **go to hell** in this way can also offend the hearer. Similarly, while to use the latter can emphasize what we are saying, it is also more likely to offend the hearer. It seems that they also need OFFENSIVE. Again, it is not clear how the lexicographers use the labels. The same goes for Sense 2 of **remind**.

There are also many cases where a pragmatic label seems necessary but is not given: for example, COBAm defines Sense 5 of **delicate** as follows:

You use **delicate** to describe a situation, problem, matter, or discussion that needs to be dealt with carefully and sensitively in order to avoid upsetting things or offending people.

Although this definition is not assigned any label, such label as POLITE-NESS may be necessary, for to use a particular expression in order to avoid offending people may generally be considered polite. There appear to be confusions between the style and the pragmatic labels in COBAm. It seems fair to say that there is some room for improvement in COBAm’s treatment of the style and pragmatic labels. (Section 4 by Kawamura)

5. **Examples**

The functions of examples in a dictionary are manifold. Landau (2001: 207–208) says, “Illustrative quotations can convey a great deal of information about collocation, variety of usage (degree of formality, humorous or
sedate context), connotation (affective implications), grammatical context (if a verb, does it take an indirect object?), and, of course, designative meaning”. Major functions of the examples in a learners’ dictionary will be as COB3 (pp. x–xi) describes: “to show the patterns that are frequently found alongside a word or phrase”, and “the typical structures in which the word or phrase is most often found”. The examples “can be used to show the characteristic phrasing round the word”, while they also “help to show the meaning of the word by showing it in use”.

In this section, we’ll compare the examples in COBAm and those in COB5. COBAm says it is “specifically designed for the American market” (Introduction: xii). Therefore, we’ll examine what changes have been made in the examples for the learners of American English. Our main focus of comparison is on the nature and the extent of new incorporations, deletions, and alterations of examples. We will examine the COBAm’s examples in comparison with those in LDOCE4 and OALD7 later in this section. The comparison will clarify the different policies of exemplifying senses among these dictionaries. It will also clarify whether COBAm’s examples show the typical patterns, grammatical constructions, and collocations of the word or phrase for the learners of American English.

### 5.1. Comparison between COBAm and COB5

In this subsection we will list numerical changes of the examples in COBAm as compared with those in COB5. The survey has been conducted on about 2% of all the examples included in COBAm and their equivalents in COB5. We will summarize the general trends with regard to deleted, altered, and newly incorporated examples in the following subsections. Table 5.1 shows how many examples have been given in COB5 and how many examples there are in COBAm, how many examples have been deleted from COB5 and how many examples have been added in COBAm, and how many examples have been changed (either partly altered or completely replaced).

Since the number of examples varies from page to page, the list gives only a sketchy representation of the changes. It is obvious, however, that COBAm shares the majority of examples with COB5 and that it has many fewer examples compared with COB5. Our sampling suggests about ten examples have been deleted per page. We suppose that the examples have been decreased to give more space for new features like Picture Dictionary, Word Webs, and Thesaurus in COBAm. The number of examples given in the front matter of the second and third editions is about 100,000 and 105,000 respectively. According to COB5, there are over 75,000 examples. COBAm does not specify the number of examples, but we estimate there are around 60,000. The number of examples has been in gradual decline since the third edition.

#### 5.1.1. Deletion

**5.1.1.1.** The most obvious reason for the deletion of examples is due to the omission of the senses for which examples are given. A large number
of senses that are marked as [(mainly)brit] in COB5 have been dropped from COBAm together with the examples. Here are some dropped examples.

1. If someone or their bank account is in credit, their bank account has money in it. [(mainly brit)] The idea that I could be charged when I'm in credit makes me very angry . . . Interest is payable on credit balances.
2. A fixture is a sports event which takes place on a particular date. [(brit)] City won this fixture 3–0 last season.

Also, a fair number of idiomatic phrases listed either as run-ons or as one of the senses in COB5 have been dropped from COBAm.

3. If you describe someone or something as a fixture in a particular place or occasion, you mean that they always seem to be there. [(BRIT)] She was a fixture in New York's nightclubs . . . The cordless kettle may now be a fixture in most kitchens.
4. If you hang around, hang about, or hang round, you stay in the same place doing nothing, usually because you are waiting for something or someone. [(INFORMAL)] He got sick of hanging around waiting for me . . . On Saturdays we hang about in the park . . . those people hanging round the streets at 6 am with nowhere to go.

5.1.1.2. When COB5 has two or more examples, the second and/or the third and the following examples have often been dropped from COBAm as illustrated in the following examples. (The examples in the parentheses are dropped.)

1. To invade a country means to enter it by force with an army. [(In autumn 1944 the Allies invaded the Italian mainland at Anzio and Salerno . . . (The Romans and the Normans came to Britain as invading armies.)]
2. You use many to indicate that you are talking about a large number of people or things. [(I don't think many people would argue with that . . . Not many films are made in Finland . . . (Many holidaymakers had avoided the worst of the delays by consulting tourist offices . . . Acting is definitely a young person's profession in many ways.)]

The reason for the omission of the second/third/etc. sentences, we suppose, is that they are less central in terms of sentence structures and collocational patterns. In addition, some of these examples contain Britishism and/or British geographical names, and things British.

5.1.2. Changes

The extent of changes varies. Minor changes have been made in the spellings, words, punctuations, etc. from British to American. [(British geographical names and things British have been changed to American geographical names and things American. Some examples have been completely replaced with new ones. Many other changes come between them.)]

5.1.2.1. Here are some examples of the minor changes (The words in parentheses in COB5 have been replaced with the underlined words in COBAm or simply dropped. Words in boldface (our modification) show the headword.): [(The characters complain ceaselessly about food lines (queues), prices and corruption. [(The Britain's) new ambassador to Lebanon has presented his credentials to the President. [(He is one of the greatest (British) players of recent times and is a credit to his profession. [(Mr. (Mr) Carter is due in Washington (London) on Monday. [(Gray (Grey) sheets flapped on the clothes line . . . an eighteenth century mansion in New Hampshire (Hampshire) US infantry (Scottish) battalion was marching down the street.)]

5.1.2.2. Replaced examples (the COB5's examples before the arrow have been replaced with the COBAm's example after the arrow): [(Joanna has finally made it to the first rank of celebrity after 25 years as an actress. —> He achieved celebrity as a sports commentator. . . . a mixture of wet sand and cement. —> Builders have trouble getting the right amount of cement into their concrete. Credit cards can be handy — they mean you do not have to carry large sums of cash. —> The book gives handy hints on looking after indoor plants.)]

5.1.2.3. Partly changed examples: [(Now he needs only two credit hours to graduate. —> After all his hard work, he now needs only two credit hours to graduate. . . . the duly elected president of the country. —> He]
is a duly elected president of the country and we're going to be giving him all the support we can. She called Amy to see if she had any idea of her son's whereabouts. As it happened, Amy had. He called Amy to see if she knew where his son was. As it happened Amy did know. ... his great powers of invention. Perhaps, with such powers of invention and mathematical ability, he will be offered a job in computers.

We can see improvements in many of the partly changed examples in COBAm. Their semantic background is given more fully and they are therefore easier to understand. Still, there are some examples in which we see no improvements like: Tom celebrated his 24th birthday two days ago. Dick celebrated his 60th birthday Monday. The reason for the replaced examples in COBAm is less obvious. We see no improvements in them, either.

5.1.3. Addition

New examples have been given where new headwords have been incorporated. Likewise, new examples have been given where new senses have been incorporated. A small number of examples have been incorporated where there had been none.

5.1.3.1. New examples for new headwords: Eva specializes in one of the most difficult techniques of basketry. Today, the student body is roughly 60 percent Black and Hispanic, 25 percent Asian, and 15 percent white. City College of New York has one of the most diversified student bodies in the nation. Jim Blachek, student council president at Sandburg High School.

5.1.3.2. New examples for new senses: Sophie turned out to be such a flake. She said she'd meet me here and instead I'm just lying around this hotel room and I'm totally bored. The government has been offering tax credits, accelerated depreciation, and other economic hanky-panky.

5.1.3.3. New examples where there were none: The dollar climbed about 30 basis points during the morning session. It was a great night and I had a massive hangover.

Most prominent changes in examples, as the comparison of COBAm and COB5 has clarified, represent adaptations to Americanism, American place names, and contextualization. The omission of a large number of examples, however, has made the dictionary less informative. We think more examples are necessary to illustrate typical structures and collocations. We will discuss this problem in the next subsection.

5.2. Comparison of COBAm with LDOCE4 and OALD7

In this subsection, we'll compare examples in three learners' dictionaries; COBAm, LDOCE4, and OALD7. The aim is to examine whether COBAm gives informative examples that show the typical patterns and grammatical structures of the headword in a readable, learner-friendly way in comparison with the other two dictionaries.

We have compared examples from two nouns, two verbs, and one conjunction for this purpose. Space doesn't allow us to deal with a larger number of examples.

5.2.1. Nouns compared

5.2.1.1. In COBAm, the first sense of door is defined as: “A door is a piece of wood, glass, or metal, which is moved to open and close the entrance to a building, room, cupboard, or vehicle”.

This canonical definition gives three pieces of information; the material (wood, glass, or metal), the function (open and close), and the location (building, room, cupboard, or vehicle) of the door. One example follows the definition. I was knocking at the front door, but there was no answer.

The example shows the combinations: “knock at the front door”, and covertly “answer (the door)”, but the most basic function of the door, namely “open the door and close the door” is not exemplified. The phrase answer the door is overtly expressed in the fifth sense: Carol answered the door as soon as I knocked. And the non-canonical phrases shuts the door in your face and slams the door in your face, defined as “refuse to talk or give you any information”, are given in the eleventh sense: Did you say anything to him or just shut the door in his face?

5.2.1.2. LDOCE4 provides much more information. First, it gives [verb–door] collocations, then [door–verb] collocations, [adjective–door]
collocations, [noun—door] collocations, etc. each followed by the example(s). (Definitions other than COBAm's have been dropped to save space.)

open/close/shut the door Could you open the door for me? | the door swung/flew/burst etc open The door flew open and Ruth stormed in.
| slam the door (= shut it loudly, usually because you are angry) | front/back/side door (= at the front, back etc of a house) Is the back door shut? | the main door (= at the front entrance) of the cathedral |

knock on/at the door Knock on the door and see if they're home. | cupboard/fridge etc door Mary slid back the closet door. | door handle/knob brass door handles

As we can see, LDOCE4 exemplifies the most basic [verb—door] combinations in the first place, then the [door—verb] combinations that display the manner in which doors open, followed by the locational information.

5.2.1.3. OALD7’s formatting is similar to that of LDOCE4’s. It gives collocational examples in boldface, as well as phrasal examples, and the sentence examples. It gives more information than COBAm but less than LDOCE4: a knock on the door ♦to open/shut/close/slam/lock/bolt the door ♦to answer the door (= to go and open it because sb has knocked on it or rung the bell) ♦the front/back door (= at the entrance at the front/back of a building) ♦the bedroom door ♦the door frame ♦a four-door saloon car ♦the fridge door ♦Shut the door! ♦Close the door behind you, please. ♦The door closed behind him.

5.2.1.4. COBS’s examples include the [open—door] combination, but it has been dropped in COBAm. We do not know the reason for the deletion. COBAm does not give any [door + verb] constructions. We believe more [verb—door] constructions, and the [door—verb] constructions should be included.

5.2.1.5. Our next comparison is money. COBAm’s first sense is defined as: Money is the coins or bank notes that you use to buy things, or the sum that you have in a bank account.

The definition gives the information on the forms (coins and notes) and the function (use (money)) of money. The examples, A a lot of the money that you pay at the movies goes back to the film distributors, and A Players should be allowed to earn money from advertising, include the combinations “a lot of money”, “pay (money)”, and “earn money”, but more common expressions like “spend money” are not included. “Make money” is given as a phrase in the sixth sense: ♦... the only part of the firm that consistently made money.

5.2.1.6. LDOCE4 offers a large number of collocations and the examples in the boxes:

| spend money | make/earn money | lose money (= not make a profit, so that a business owes more than it earns) | cost money / cost a lot of money | save money | lend/borrow/owe money | waste money (on sth) | be a waste of money | charge (sb) money | raise money | pay money (for sth) | give sb their money back / refund sb’s money |
| a sum / an amount of money | get/earn good money (= be paid good wages) |

Don't spend all your money on the first day of your holiday! She doesn't earn very much money. He's working for a finance company now, and making loads of money. At last the business is starting to make money. The company is losing money and may have to close down. The repairs will cost quite a lot of money. We're not going on holiday this year because we're trying to save money. Could you lend me some money? I don't want to borrow money from the bank unless I really have to. They charge huge amounts of money for their services. We're trying to raise money to help children with cancer. If you are not completely satisfied with our products, we will give you your money back. He was left a large sum of money. You can earn good money as a computer programmer.
We prefer *LDOCE4*’s presentation of collocations and examples to *COBAm*’s definition-and-the-examples format. It lists a wider range of [verb–money] combinations as collocations. It contains the most basic verbs related to “income” and “expenditure” such as “earn/make money”, and “spend/cost money”. Moreover, many more [verb–money] combinations are exemplified. Since money is one of the most frequent topics in the everyday situations and there are various money related expressions, learners will naturally expect to find a large number of money related expressions in the dictionary.

5.2.1.7. *OALD7*’s presentation is less informative than *LDOCE4*’s, but it gives more information as collocations and examples than does *COBAm*:

- to borrow / save / spend / earn money
- *How much money is there in my account?*  
  *The money is much better in my new job.*  
  *If the item is not satisfactory, you will get your money back.*  
  *We’ll need to raise more money (= collect or borrow it) next year.*  
- *Can you lend me some money until tomorrow?*  
  *Be careful with that — it cost a lot of money.*

The expression “make money” is given as an idiom: *The movie should make money.*  
*There’s money to be made from tourism.*

5.2.1.8. Learners will not find it difficult to grasp the meanings of nouns like *door* and *money*. If they consult a dictionary for these nouns, they will be more likely to look for words that combine with them. *COBAm* might better provide learners with many more typical patterns in which these words are used.

5.2.2. Verbs compared

5.2.2.1. Our next comparison is the verb *eat*. The first sense of *eat* in *COBAm* is defined as:  
1 V–I/V–I When you *eat* something, you put it into your mouth, chew it, and swallow it.  
2 *She was eating a sandwich*.  
3 *I ate slowly and without speaking.*

The second sense is defined as:  
2 V–I If you *eat* sensibly or healthily, you eat food that is good for you.  
4 . . . a *campaign to persuade people to eat more healthily*.

In addition, the phrases “*eat too much*”; “*eat properly*”, “*eat well*”; “*eat alone*”, “*eat together*”; “*want something to eat*”; “*eat and drink*”, “*eat and sleep*” are given as Word Partnership in the box.

The examples are easy to understand and show the grammatical structures of the verb, but only three examples are too few for the first and second senses. Aren’t there any more expressions that should be included in the dictionary?

5.2.2.2. *LDOCE4* lists the collocations, phrases, and idioms in the box below. The paraphrases in the parentheses will help learners. Thirteen examples are given in the following box.

<table>
<thead>
<tr>
<th>Table 5.3</th>
</tr>
</thead>
</table>
| *eat well/healthily/sensibly* (= eat food that will keep you healthy)  
*eat right* *AmE* (= eat food that will keep you healthy)  
*eat properly*  
*eat healthily*  
*eat right* and *get plenty of sleep*  
*Would you like something to eat?*  
*She can eat like a horse and never put on weight. *  
*I couldn’t eat another thing/bite*  
*I couldn’t eat another thing.* |

*Felix chatted cheerfully as he ate.*  
*An small girl was eating an ice cream.*  
*We had plenty to eat and drink.  
It’s important to *eat healthily* when you are pregnant.  
*I exercise and eat right and get plenty of sleep.*  
*Would you like something to eat?*  
*She can eat like a horse and never put on weight.  
*We stopped at McDonald’s to get a bite to eat.*  
*Good eating habits are the best way of preventing infection.*  
*ready-to-eat foods such as deli meats and cheeses*  
*More cake?’ No thanks, I couldn’t eat another thing.’*  
*No chicken for me. I don’t eat meat (= I never eat meat).  
*Does Rob eat fish?* |

*LDOCE4* gives many more collocations and examples that show the [verb–adverb] combinations including *eat right* *AmE*. This particular expression is absent in *COBAm*, the dictionary that targets the American market. Instead, it lists “*eat properly*” which is labeled as *BrE* in *LDOCE4*.

We believe the last two examples in *LDOCE4* are worth inclusion in *COBAm* because they are used to express eating habits and are often heard in everyday situations.
5.2.2.3. OALD7 gives fewer examples than LDOCE4 but it gives more examples than COBAm. Here are OALD7's examples: [V] I was too nervous to eat. ◇ She doesn't eat sensibly (= doesn't eat food that is good for her). ◇ [VN] I don't eat meat. ◇ Would you like something to eat? ◇ I couldn't eat another thing (= I have had enough food). OALD7 gives the expression “eat like a horse” as an idiom with the example: She may be thin, but she eats like a horse.

5.2.2.4. As in the case of door and money, the examples for this verb in COBAm are too pared down. High frequency words should be provided with more examples.

5.2.2.5. We will now compare another verb see. Followings are the fifth definition and the examples for see in COBAm: 5 V–T If you see what someone means or see why something happened, you understand what they mean or understand why it happened. [no cont, no passive] ◇ Oh, I see what you're saying. ◇ I really don't see any reason for changing it.

COBAm gives three other use of see in the sense “to understand” as PHRASE and CONVENTION: 6 CONVENTION ◇ “He came home in my car.” — “I see.” 7 PHRASE ◇ He was hopeful that by sitting together they could both see sense and live as good neighbors. 8 CONVENTION ◇ Well, you see, you shouldn't really feel that way about it.

We can see from the definition 5 that see in this sense typically takes [see–what] and [see–why] constructions. The first example illustrates this construction. But we know see in this sense has many more everyday expressions such as those given in LDOCE4.

5.2.2.6. LDOCE4's examples: see why/what/how etc I can't see why he's so upset. ◇ I see what you mean (= I understand what you are saying). ◇ 'He lives here but works in London during the week.' 'Oh, I see (= I understand).’ ◇ You see, the thing is, I'm really busy right now (= used when you are explaining something). ◇ You mix the flour and eggs like this, see (= used to check that someone is listening and understands). ◇ I can't see the point of (= I do not understand the reason for) spending so much money on a car. ◇ Do you see the point I'm making (= do you understand what I'm trying to say)? ◇ The other officers laughed but Nichols couldn't see the joke. ◇ see reason/sense (= realize that you are wrong or doing something stupid) I just can't get her to see reason!

5.2.2.7. OALD7 also gives a variety of everyday expressions: [V] ‘It opens like this.' ‘Oh, I see.' ◇ [VN] He didn't see the joke. ◇ I don't think she saw the point of the story. ◇ I can see both sides of the argument. ◇ Make Lydia see reason (= be sensible), will you? ◇ [V (that)] Can't you see (that) he's taking advantage of you? ◇ I don't see that it matters what Josh thinks. ◇ [V wh–] ‘It's broken.' ‘Oh yes, I see what you mean.' ◇ Can we go swimming?’ ‘I don't see why not (= yes, you can).’ ◇ [VN to inf] The government not only has to do something, it must be seen to be doing something (= people must be aware that it is doing sth).

OALD7’s examples include not only [see–wh] constructions but also [see–(that)] constructions.

5.2.2.8. Here again, COBAm uses only a small number of examples to illustrate the definitions and does not illustrate a sufficient number of common structures and expressions. We feel many more examples should be added for the learners.

5.2.3. Conjunction compared

Our last comparison is the conjunction whether. Unlike the content words like door, money, eat, and see whose meaning is fairly transparent, the function words like whether often depend greatly on the use of language to supplement their meaning.

5.2.3.1. The definitions and the examples of whether in COBAm are:

1. You use whether when you are talking about a choice or doubt between two or more alternatives. ◇ To this day, it's unclear whether he shot himself or was murdered. ◇ Whether it turns out to be a good idea or a bad idea, we'll find out. 2. You use whether to say that something is true in all of the circumstances that you mention. ◇ This happens whether the children are in two-parent or one-parent families. ◇ Whether they say it aloud or not, most men expect their wives to be faithful.

The definitions and examples lack two important uses of this conjunction: Like if, whether is used as an interrogative subordinator, in addition, it takes the [whether + to–inf] construction. Both of these uses are exemplified in LDOCE4 and OALD7.
5.2.3.2. LDOCE4 divides the senses in two like COBAm, and it provides the examples for the [whether + to-inf] construction, as well: Maurice asked me whether I needed any help. | There were times when I wondered whether or not we would get there. | *whether to do sth* She was uncertain whether to stay or leave. | I didn’t know whether to believe him or not. | The question arose as to whether this behaviour was unfulfilling.

2 It seemed to me that she was in trouble whether Mahoney lived or died. | Look, Kate, I’m calling the doctor, whether you like it or not. | Poor farmers, whether owners or tenants, will be worst affected.

5.2.3.3. OALD7 provides twice as many examples as COBAm for the two senses: 1 He seemed undecided whether to go or stay. | It remains to be seen whether or not this idea can be put into practice. | I asked him whether he had done it all himself or whether someone had helped him. | I’ll see whether she’s at home (= or not at home). | It’s doubtful whether there’ll be any seats left. 2 You are entitled to a free gift whether you accept our offer of insurance or not. | I’m going whether you like it or not. | *Whether or not* we’re successful, we can be sure that we did our best.

5.2.3.4. We don’t know why COBAm doesn’t give examples of these two important uses of *whether*. Are they less frequently found in the corpus? We believe these two uses should be incorporated in the dictionary whether or not they appear frequently in the corpus. We believe a pedagogical dictionary should not depend solely on the frequency to decide which constructions and examples to include.

We believe that the use of pronouns without clear contexts is undesirable. The use of personal pronouns like “he”, “she” and “they” may be inevitable to save space. Learners will take them as the conventions that stand for a male/female person or persons, whoever they may be. However, context dependent pronouns like “this” in COBAm: “This happens whether the children are in two-parent or one-parent families” should be avoided. Although it is not the focus of the example, the meaning of the sentence is opaque without the context that specifies the pronoun “this”.

We have compared only two nouns, two verbs and one conjunction, so the comparison may not be sufficient to draw any conclusions. Even so, we think COBAm’s examples for core entries are insufficient compared with other learners’ dictionaries because we can offer a lot of evidence to prove this from other lexical categories. One reason for fewer examples may be that COBAm provides examples evenly for every sense that it believes deserves exemplification instead of giving more examples for core entries. For example, COBAm gives examples for Dolby [trademark], hard cash, life history, moviegoer (cinemagoer), etc. | ... a cassette deck equipped with Dolby noise reduction. | There is no confusion about what the real dividend is since the payment comes in hard cash. | Some people give you their life history without much prompting. | What is it about Tom Hanks that moviegoers find as appealing? LDOCE4 and OALD7 do not offer any examples for these entries. Some learners may find these examples very helpful, but most learners can surely dispense with them and want more examples for common, everyday words and phrases instead.

5.3. Semantic contents

Finally, we focus our attention on the issue of the quality of examples. We will discuss this issue in terms of the semantic content of examples. Béjoint (1994, 2000: 135) states that “The choice of the semantic content of examples is relatively free, since examples are mainly meant to illustrate syntactic behaviour or to provide additional semantic information”. This statement is meant for general purpose dictionaries and not for learners’ dictionaries. But it holds true for learners’ dictionaries as well.

There seem to be no generally accepted norms as to what kind of semantic information a learners’ dictionary should offer. So, we will examine what kind of semantic contents are given in COBAm and whether they are appropriate for learners.

5.3.1. Semantic contents of COBAm, OALD7, and LDOCE4

5.3.1.1. COBAm’s examples all seem fine in that they do not put forward any strange examples such as are described in Béjoint. Many examples depict familiar, everyday situations that are easy for learners to understand (Words in boldface (our modification) show the headword): Their French has improved enormously. Thousands have lost their jobs. I must thank...
you for being so kind to me. ❑Is there life on Mars? ❑He studied his map, trying to memorize the way to Rose's street. ❑He was going to college at night, in order to become an accountant.

However, there are examples whose topics are related to political, diplomatic, economic, social, military, etc. issues which probably have been taken from newspaper articles. These topics tend to be specific to occasions in the past and therefore less familiar to learners. Here are some typical ones:

❑He alleged that he was verbally abused by other soldiers.
❑The Russian foreign minister yesterday canceled his visit to Washington.
❑Two American soldiers were among the dead. ❑The party fears the equation between higher spending and higher taxes. ❑Consumers have a wide array of choices and price competition is fierce. ❑. . . a global ban on nuclear testing. ❑The general divided his time between his shabby offices and his home in Hampstead.

We think these examples are well construed to exemplify the definitions. But we feel there should be better examples with more learner-friendly semantic contents.

5.3.1.2. We'll compare the examples of the verb promise in the three dictionaries to see what kinds of topics are chosen. We can see that OALD7 and LDOCE4 cite more familiar examples that are easy for learners to understand. COBAm's examples, although they include easy ones, seem more appropriate for general readers rather than for learners.

COBAm's examples:
❑The post office has promised to resume first class mail delivery to the area on Friday.
❑He had promised that the rich and privileged would no longer get preferential treatment.
❑Promise me you will not waste your time.
❑I'll call you back, I promise.
❑In 1920 the great powers promised them an independent state.
❑While it will be fun, the seminar also promises to be most instructive.

We feel that some of these examples in COBAm are too crude for learners. They might have been chosen directly from newspaper articles or the like without modification. Their phraseology and the sentence structures are generally more difficult than those in OALD7 and LDOCE4. As mentioned above, topics related to politics, economics, social issues, etc. that were reported in the newspaper many years ago are unlikely to be interesting to the learners.

OALD7's examples: The college principal promised to look into the matter. ❑'Promise not to tell anyone!' 'I promise.' ❑They arrived at 7.30 as they had promised. ❑The government has promised a full investigation into the disaster. ❑I'll see what I can do but I can't promise anything.
❑The brochure promised (that) the local food would be superb. ❑You promised me (that) you'd be home early tonight. ❑He promised the money to his grandchildren. ❑I've promised myself some fun when the exams are over. ❑I'll be back soon,' she promised.
❑It promises to be an exciting few days. ❑There were dark clouds overhead promising rain. ❑I can promise you, you'll have a wonderful time. ❑If you don't take my advice, you'll regret it, I promise you.

These examples in OALD7 are likely to be much easier for the learners to read because the topics are more familiar to them: dialogues between friends and family, and topics relating to the everyday situations. Most words used in these examples are listed in The Oxford 3000.

LDOCE4's examples:
❑Last night the headmaster promised a full investigation. ❑She's promised to do all she can to help. ❑Hurry up — we promised we wouldn't be late. ❑You promised me the car would be ready on Monday. ❑'Promise me you won't do anything stupid.' 'I promise.' ❑I've promised that book to Ian, I'm afraid. ❑The company promised us a bonus this year. ❑'Promise?' 'Yes! Don't worry.' ❑He reappeared two hours later, as promised. ❑Tonight's meeting promises to be a difficult one. ❑dark clouds promising showers later ❑I'll try my best to get tickets, but I can't promise anything. ❑I promise you, it does work!

LDOCE4's examples are similar to those of OALD7. They include dialogues and familiar everyday topics in easy phrasing.

5.3.2. Genuine vs. adapted examples
5.3.2.1. The comparison above and the comparisons in 5.2 seem to suggest that the semantic contents in the three dictionaries are deeply related to the sources of examples and how they are presented. Two different views seem to exist when dictionaries present examples chosen from the corpora.
Genuine examples chosen from the corpus (The Bank of English) have been one of the major features of the COBUILD dictionaries from the very beginning.

...the examples in this dictionary have a new status and do a different job from examples in the conventional tradition. These examples are taken from actual texts wherever possible. ...in recent years it has become a habit to make up examples to illustrate a particular usage. These examples have no authority ... (COB1: (p. xv))

Although examples in COBUILD dictionaries are not exact copies from the citation files, they seem to have been given in a form that is as close to the original text as possible with minimal changes. The source materials for many of the citations seem to include articles from newspapers and/or other journalistic texts. Since these source materials are deeply related to a knowledge of the real world, learners with limited knowledge and experiences will find the examples rather difficult to understand. As we have mentioned above, examples cited from past newspaper articles may not be very interesting for many learners.

5.3.2.2. OALD7 expresses quite a contrary view:

A corpus is an abundant source of samples of English usage. If they are to function as examples, however, then we need to ask just what it is they are intended to be examples of. ... Samples of the language, isolated from their natural context of use, will not normally exemplify word meanings, but will simply show one instance, among innumerable others, of the word's actual occurrence, which, in itself, is of little if any help to the learner. ... In view of these obvious difficulties, rather than insist on the use of authentic samples, it would seem to make more sense for a learner's dictionary to follow the quite different principle of pedagogic appropriateness and to provide, as OALD does, examples that are designed for the express purpose of exemplification. This does not mean that any less account is taken of the findings of corpus descriptions, but only that they are taken into account in ways which make them more accessible and relevant to learners. (OALD7 (p. viii))

The examples in a learners' dictionary should not be unnatural English that will not be used outside the dictionary. At the same time, they should be simple, easy English that learners can understand within their scope of interest and knowledge. Examples in descriptive dictionaries that record the language facts and those in pedagogical dictionaries that help learners acquire the language should necessarily be different.

COBAm's topics in its examples have been chosen from various aspects of our world. They include global and local politics, the international and domestic economy, social affairs, religion, accidents, conflicts, etc. It is true that these topics are reported on TV and in newspapers every day. Naturally, the corpus from which COBAm's examples are chosen may have an extensive stock of data on these topics. However, examples chosen from the corpus according to frequency are unlikely to be very helpful for learners. They sometimes contain topics that are unfamiliar to them. The sentence structures in actual use can be very complicated and the words and phrases can be rather difficult. If good examples cannot be found in the citation files, lexicographers would do well to make more radical adaptations as LDOCE4 and OALD7 seem to have done.

(Section 5 by Masuda)

6. Vocabulary Builders

Vocabulary Builders consist of the following: Picture Dictionary boxes, Word Webs, Word Partnerships, Word Links, Thesaurus entries and Usage notes. They are innovation found only in COBAm. The features of Vocabulary Builders will be discussed in the subsections below.

6.1. Picture Dictionary boxes

Picture Dictionary boxes (PDs) are multi-colored illustrations intended to visually describe vocabulary or concepts in COBAm. According to the front matter, the words described in Picture Dictionary boxes are selected based on "their usefulness in an academic setting, frequently showing a concept or process that benefits from a visual presentation" (p. vii).

The number of PDs in the printed version of COBAm does not coincide with that of the PDs on the CD-ROM; there are fifty PDs in the
printed version, but two of them are missing from the CD-ROM. The PD for **house** is not found on the CD-ROM, while the PD for **farm** on the CD-ROM is identical with the Word Web (WW) for **barn** in the printed version.

Forty-eight out of the fifty PDs describe nouns. The remaining two are for the verbs **cook** and **cut**. The words described in PDs do not necessarily seem to be those which are useful in an academic setting as it was stated in the front matter but appear to be more general words which may also be of use in our daily life such as **bread**, **dessert** and **egg**. Thus it can be said that the PDs are not only intended to give academic information concerning the headwords; rather they are intended to give encyclopedic information on the selected headwords.

PDs are inserted in the body of the dictionary. This seems to be unusual since dictionaries are normally concerned about the space. The approach in **COBAm** gives easier access to PDs since they are provided near their headwords.

The PD data are combined with the WW data on the CD-ROM. A click on the “Word Web & Picture Dictionary” tab brings out a complete list of the PDs and WWs on the left-hand side of the screen. A PD is shown in full size by clicking on the word in the list.

### 6.1.1 Types of illustrations

Illustrations in PDs are similar to those in Duden picture dictionaries, consisting of “the grouping of drawings and the detailed labeling of the various parts in each illustration” (Landau 2001: 147). For example, for **bed**, the names of the various parts which compose a bed such as a **bedspread**, a **canopy**, a **pillow**, and so on are labeled in the illustration of a bed.

According to the classification of dictionary illustrations given in Svensén (1993: 172–176), the illustrations in PDs can be classified into the following three types:

1) Several objects of the same class e.g. **dish**
2) Environment with typical objects e.g. **golf**
3) Basic objects and concepts in a subject field e.g. **family**

For example, for the word **dish**, various kinds of dishes such as a **bowl**, **plates**, a **platter**, etc. are depicted to show the variation of the shapes and size of containers that are considered as dish as one group. The PD for **golf** shows a **golfer** and a **caddie** on a golf course. The golf gear used by the golfer and the items used on the golf course are labeled. The PD for **family** is a family tree with all the members in the tree labeled based on their relationship to one member in the family tree.

Photographs are not used in PDs, while they are used along with pictures in WWs. This has three advantages. First, it is easier to paint a picture which comprises a set of items associated to a headword than to take a photograph of it. For example, in the PD for **desert**, the picture depicts a **buzzard**, a **lizard**, a **scorpion** and a **snake** with some plants that are found in a desert situated in a typical desert-like landscape. It may not be easy to take a photograph of this scene; for different kinds of birds, animals and plants are not expected to appear in the same place at the same time.

Second, it is easy to deform a part or parts of a picture for the users’
convenience. For example, the stationery goods on the desk are magnified in the PD for office so that the users can clearly see the small items such as a highlighter, paper clips and sticky notes on the desk.

Lastly, certain items are almost impossible to be photographed. An example is the PD for core, in which the inner part of the globe is shown in the picture.

6.1.2. Words chosen for Picture Dictionary boxes

As mentioned above, the words chosen for PDs are not always “chosen for their usefulness in an academic setting” (p. vii). What kinds of vocabulary merit visual presentation? This subsection focuses on this question and offers three reasons.

First, words and concepts that have culture-specific information are worth describing visually, because pictures can clearly show a typical example of a reference which a word denotes. For example, the PD for egg illustrates several different ways of cooking eggs along with the names for each dish. Cooking is closely tied with culture; a familiar dish for the Americans may be quite foreign to people who have different cultural backgrounds.
Second, pictures can help the typical users of \textit{COBAm}, i.e. EFL learners. They often have similar names for concepts or objects in their native language. Thus a picture may help them associate the reference with its name in the target language. Even if they lack the reference in their native language, they would be able to identify one in the future (Landau 2001: 146, Svensén 1993: 168f).

Third, there are certain concepts and objects which can be grasped more easily by looking at pictures than by reading through complicated definitions. Typical examples are PDs for \textit{area}, \textit{color}, \textit{shape}, and \textit{volume} for concepts, and \textit{cloud}, \textit{globe}, \textit{landform}, and \textit{river} for objects.

6.1.3. Encyclopedic redundancy

There are a couple of PDs that, in a sense, seem to show encyclopedic redundancy. For example, \textit{COBS} gives only four examples of the numerals in the entry for \textit{Roman numerals}. On the contrary, \textit{COBAm} shows an extensive list of the representative Roman numerals in a table. If a user thinks it is enough to see several examples to get the idea of what the Roman numerals are, then the examples in \textit{COBS} are sufficient. If, however, the user wants to know the rule of how to write numbers in Roman numerals, then he/she benefits from the table in \textit{COBAm}. The same can be said with the PD for \textit{sign language}. Whether the information is redundant or not depends on the user’s needs.
6.2. Word Webs

A Word Web is a box which consists of one or more than one pictures or photographs which depict a typical feature of a headword and a passage which gives encyclopedic information on the headword. The passage contains several key words which are semantically related to the featured headword. They are printed in bold if they are available at the entries in COBAm. According to the explanation given in the front matter, the users of the dictionary are encouraged to read through the WWs to get "deeper understanding of the language and concepts" (p. viii).

The number of WWs in the printed version of COBAm (248) does not coincide with that of the WWs on the CD-ROM (224). To detail the discrepancy, the WWs which are in the printed version but are not included on the CD-ROM are for the following headwords: astronomer, bird, country, drum, echo, eclipse, English, history, meat, medicine, population, pottery, radio, refrigerator, science, solar system, sound, sun, tree, vote, wave and year. The WW for tobacco is not in the printed version but is on the CD-ROM, while the WW for crash in the printed version is included as the WW for break on the CD-ROM.

Besides the discrepancy in number, there are a couple of differences between the WWs in the printed version and those on the CD-ROM. There are five WWs whose pictures and photographs in the printed version are not the same as the ones used on the CD-ROM. They are the WWs for the following headwords: flag, meal, skin, sweat and writing. In addition to them, the PD for barn in the printed version is included as a PD for farm on the CD-ROM.

6.2.1. Key words

The words highlighted in bold in WW passages are the key words that are thematically related to the featured headwords. They help the users learn a headword by forming a lexical network in their minds with the featured headword in the center. The name Word Web may have originated from the image of the lexical network that would look like a spider's web.

The key words are usually available at the entries. An asterisk is attached if the definition is not available in COBAm. For example, the meaning of the word signature in the WW for book is not available in COBAm. In this case, signature denotes a group of pages. Therefore, an asterisk is attached to it and the definition is given at the end of the passage so that the user does not have to look it up in another dictionary.

The number of key words in a WW varies. There are 19 key words in 11 lines in the WW for bank, while there are only 3 key words in 5 lines in the WW for zero.

6.2.2. Pictures and photographs

There are usually one or two pictures or photographs in a WW. There are as many as five pictures in the WW for amphibian but this is unusual. The picture or photograph often shows a typical example of the headword. However, it is sometimes difficult to describe arbitrary concepts in a picture or a photograph. The following is the illustration in the WW for genre.

![Figure 6.8 Illustration in the WW for genre](image)

Scientific terms are often provided with WW boxes with drawings in a Picture-Dictionary style. That is, the illustration is a drawing with detailed labeling of its parts. Some of the typical examples are the illustrations in the WWs for engine and respiratory system.

The pictures and photographs in the WWs not only function as eye-catchers, but they also show the items that are related to the headword.
For example, in the WW for painting, items used for oil painting such as a canvas, paintbrushes, and a palette are all shown in the photograph.

Pictures in WWs are not only drawings. Sometimes diagrams are used to convey the encyclopedic information more clearly to the users. For instance, graphs are used in the WWs for interest rate and population, while maps are used in the WWs for country and empire. There are also chronologies in the WWs for history and medicine. The content of the passage is sometimes illustrated in the picture as can be seen in the WWs for eclipse and photosynthesis.

Only pictures can be used to show typical examples of the words which denote objects or concepts that do not exist as entities in the real world.

The WWs for fantasy and myth are typical examples. Pictures are also convenient to show a range of items that are included in a headword. For example, the picture of orchestra includes various kinds of musical instruments that are usually played in an orchestra.

However, photographs are also used in abundance instead of pictures in WWs. For example, meteorological phenomena such as a hurricane and lightning are shown in photographs. Celestial bodies such as the earth and the moon are also shown using photographs. A portrait of an inventor named Thomas Midgley, Jr. in the WW for inventor is the only photograph which portraits a particular person in COBAm.

6.2.3. Cross-references to WWs

There are cross-references to WWs in COBAm. The cross-references to the WW boxes help the users get a deeper knowledge concerning the headwords. For example, there is a cross-reference to the WW at the end of the entry for the noun can. By reading through the passage in the WW, we can learn the history of how the technique for keeping food fresh in a can was developed in the course of time along with a range of key words related to the headword.
Another example is the cross reference for GPS. At the end of the entry for GPS, there is not only a cross-reference to the headword navigation but also to the WW for GPS. Therefore, by referring to the WW, the user can not only get an image of what a GPS looks like through a photograph but also can obtain information that is not given in the entry.

6.3. Word Partnerships

Word-Partnership (WP) boxes show frequent collocations of selected headwords. It is said in the front matter that the headwords with WP boxes are chosen based on frequency in the Bank of English (p. viii). There are 1,107 WP boxes in COBAm.

There is always a phrase “Use the X with:” on the first line in the box. Words that often collocate with the headword are listed in the box. They are classified according to their part of speech. There is a number in a blue square at the end of each list of collocations. It indicates the definition number of the headword. The collocations are listed according to the definition numbers in ascending order.

The introduction of other COBUILD dictionaries was to show the collocations of a word in their definition sentences. However, only a small number of collocations — one or two collocations per definition — can be illustrated by this method. The introduction of WP boxes enabled COBAm to provide the users with a larger number of collocations per headword.

6.4. Word Links

Word-Link (WL) boxes explain the meaning of prefixes, suffixes and roots that are frequently used in English. There are 493 kinds of affixes and roots that are explained in WL boxes. Each box deals with one affix or root and gives it a simple definition and three example words. The affixes or roots in the example words are highlighted in blue. It is said in the front matter that the three examples are to encourage learners “to look up these words to further solidify understanding” (p. ix).

It is generally accepted that the knowledge of word formation helps learners to increase their vocabulary in their target language, as stated in Nation (2001):

A knowledge of affixes and roots has two values for a learner of English: it can be used to help the learning of unfamiliar words by relating these words to known words or to known prefixes and suffixes, and it can be used as a way of checking whether an unfamiliar word has been successfully guessed from context.

(Nation 2001: 264)

The WL boxes may have been introduced in COBAm as an innovation to help learners increase their vocabulary. However, the WL box is not a new idea in learner’s dictionaries. For example, in Kenkyusha’s Lighthouse English-Japanese Dictionary (henceforth Lighthouse), which is one of the English-Japanese dictionaries whose target users are mainly high school students in Japan, a similar box has been used since its second edition in 1990. The box is called “Tango no Kioku”, or “memorization of words”, which gives the knowledge of word formation to the users by explaining the meaning of affixes and roots in Japanese.

The concept of the box resembles that of the WL box. However, the number of the example words given in a “Tango no Kioku” box exceeds that of the example words in a WL box. For example, the WL box for the root port in COBAm gives only the following three example words: export, import and portable. Whereas, the “Tango no Kioku” box gives the following eight example words with the original meaning and the present meaning of each word in Japanese: export, import, report, sport, support, transport, portable and porter. Therefore, it can be said that “Tango no Kioku” boxes in Lighthouse are richer in information than WL boxes in COBAm.

6.5. Thesaurus entries

Synonyms and antonyms of selected entry words are listed in Thesaurus entries. There is only one Thesaurus entry for one headword even though the headword has more than one part of speech. COBAm says in its front matter that Thesaurus entries are attached to high frequency headwords (p. ix). In fact, 202 out of the 715 Thesaurus entries are for the headwords that have more than three frequency diamonds.

A Thesaurus entry is usually placed at the end of an entry. The
headword is printed in blue on the first line in the box. The synonyms and antonyms are listed in alphabetical order after the message "Also look up:". The list of synonyms always comes first, followed by the label (ant.), then the list of antonyms.

An entry in COBAm is not divided but unified even though a headword has more than one part of speech. This is because the dictionary is compiled based on the "One Word One Entry" system. If synonyms and/or antonyms are listed for more than one part of speech in a Thesaurus entry, the part of speech of the synonyms and/or antonyms in the list is indicated at the beginning of each list using abbreviations printed in blue. A number in a black circle indicates the part of speech of the headword. A number in a blue square indicates the definition number of the headword.

The number of the synonyms listed in a Thesaurus entry varies. It is said in the front matter that an effort is made in listing more synonyms than antonyms (p. ix). In fact, there are more synonyms than antonyms listed in all of the Thesaurus entries. Sometimes there are no antonyms listed in a Thesaurus entry, whereas there are at least two synonyms in each Thesaurus entry. The number of the synonyms and/or antonyms of a headword given in COB5 is usually small — often one or two — compared to COBAm owing to the lack of space, since they are listed in the Extra Columns. The introduction of the Thesaurus entry enabled COBAm to list more synonyms and antonyms for a headword.

There are also Thesaurus Boxes in LAAD2. Whereas only synonyms and antonyms of a headword are listed in COBAm, synonyms and antonyms are provided with definition in LAAD2, enabling the users to see the subtle differences of meaning and uses of each synonym or antonym. For example, moral has a Thesaurus entry in both COBAm and LAAD2. The following four words are given as its synonyms in COBAm: ideology, philosophy, principle and standard. In LAAD2, the synonyms are: ethics, standards, values, principles, scruples and morality. They are not only defined but also their use is illustrated with example sentences in LAAD2.

If synonyms and antonyms are listed without definitions or example sentences in a Thesaurus entry as they are in COBAm, it is presumed that only a highly motivated learner may look up all the words listed in a Thesaurus entry to know their differences in meaning and use. If, however, the synonyms and antonyms are accompanied by definitions and example sentences as they are in LAAD2, then a user can further learn the differences in the meaning and use among the listed words without taking the trouble of looking them up. Thus, it may be fair to judge that Thesaurus Boxes in LAAD2 are more user-friendly than Thesaurus entries in COBAm.

6.6. Usage

According to COBAm, "'Usage' notes explain shades of meaning, clarify cultural differences, and highlight important grammatical information" (p. ix). The total number of the notes is 96: 58 for meaning, 5 for cultural differences, and 33 for grammatical information. Each note basically consists of the heading, followed by brief explanation and examples.

Although COBAm's introduction of these notes will be welcome, those for grammar especially are sometimes too simple or too focused on a particular fact to capture the essential information on the word(s) or expression(s) explained. For example, the note on worst and worse runs as follows:

Worst and worse sound very similar. You should avoid substituting one for the other in various expressions.

The note directs users' attention to the fact that the pronunciations of the two words are similar, and advises users not to confuse them. The point the lexicographers make here is reasonable, but there are many words which can be pronounced exactly the same. If the two words are often confused only because of their similar pronunciations, there will be far better candidates. The note would not really make sense without mentioning that worse and worst are respectively comparative and superlative of bad.

It is also to be mentioned that some of the notes seem to be targeted at beginners or intermediate learners rather than advanced learners. For example, a note is given to aunt and ant which advises the user not to confuse them because their pronunciations are the same. There are also
notes which simply point out that particular words are often confused and show how to use them properly: than and then. As their examples of wrong usage often contain names foreign to English such as Xian-li (imply and infer) and Phailin (tear and break), they may have been taken from writings by foreign learners. There is a possibility that the notes were selected based on error analysis of learners' uses of particular words.

It is the COBUILD tradition to make the best of data from a large-scale corpus, which might affect their selection of words for these notes. However, their selection of items and ways of formatting explanations need to be reconsidered in terms of how useful their notes are to their target users. 

(Section 6.6 by Kawamura)

7. Functions and searchability of the accompanying CD-ROMs

This section deals with the functions and searchability of the CD-ROM versions (COBS-CD and COBAm-CD) compared to the printed ones (COBS and COBAm). We will explain the differences between the printed and the CD-ROM versions, as well as between the two CDs. Our main concern is what advantages and disadvantages the CDs have. We will argue that COBS-CD is user-friendly in many respects while COBAm-CD has much to be done to be a good reference work.

7.1. Contents and interface

First of all, we will describe what contents and functions are available in the CD-ROM versions.

7.1.1. COBS-CD

The application window (see Figure 7.1) mainly consists of the menu bar at the top, the search area on the left, and the main section for displaying the entry for the selected headword. Below the search box in the left area is a list of all the headwords and phrasal verbs in the dictionary under the tab labeled “Index”. This section changes into the area labeled “Result” where the search results are displayed when a search key is entered into the search box and the “OK” button is clicked or the “Enter” key is pressed.

![Figure 7.1 Initial screen of COBS-CD](image)

The two large buttons (Dictionary and WordBank) located in the menu bar at the top of the window allow us to choose between the following reference works:

- Dictionary: almost the same reference work as the printed version of COBS
- WordBank: a five-million-word sample of the texts from the Bank of English, a large corpus of English

COBS-CD also offers the pop-up version of COBS and “MyDictionary”, a user-defined dictionary, neither of which will be dealt with here.

The interface has the following useful features:

- The application window can be enlarged, and the main entry section can be narrowed or widened.
- The “Full/Comp” button at the top right-hand corner switches the screen between the full view of the entry and the compact view where
IPA transcriptions, the speaker icons for playing back recorded pronunciations, and examples are not displayed.

• Any part of the text can be copied to the clipboard. (COB4-CD allows us to copy either a word in the text or the whole entry.)

7.1.2. COBAm-CD

The application window (see Figure 7.2) has six different modes indicated by the tabs at the top. When the “Dictionary” tab is selected, the search area is available on the left where a list of all the headwords beginning with the character(s) typed in is displayed, and the main section is on the right for displaying the entry for the selected headword.

Figure 7.2 Initial screen of COBAm-CD

The six tabs allow us to choose between the following reference modes:

• Dictionary: We can search through content that is almost the same as the printed version of COBAm.
• My Words: This is a user-defined dictionary where headwords, parts of speech, definitions, and examples can be registered.
• Word Web & Picture Dictionary: Encyclopedic notes (Word Webs) and pictorial illustrations (Picture Dictionaries), summing up to about 300 items, can be searched for.

• Bookmarks: This offers a bookmarking function by which users can bookmark entries.
• Activity Guide: Users can learn how to use the dictionary through looking-up activities. This portion is based on the same contents in the front matter in the printed version.
• Grammar: This is an explanation of grammar using pictorial illustrations and animations (the flow of the time and focused points therein are marked for the grammatical items concerning tense), based on the same contents in the back matter in the printed version.

Dictionary and Word Web & Picture Dictionary are dealt with in this section.

We find inconvenient the following characteristics in the interface of COBAm-CD:

• The application window cannot be enlarged, nor can the search and entry areas. The amount of information that can be displayed is very small.
• The Mini-View button available in the Dictionary mode just makes the window small; this is not a so-called pop-up dictionary. In this view, additional information such as inflections in the entry are omitted, and double-clicking on a word does not execute a new search.
• More than one line cannot be copied at one time.

7.2. Dictionary contents and layout

In addition to the differences in the interfaces of the reference works, the contents of the CD-ROM versions also differ from their printed counterparts. We will see which information is and is not available in the CD-ROM versions below.

7.2.1. COB5-CD

The dictionary contents are displayed mostly in the same way as the printed version, but with the following differences:

• IPA transcriptions can be shown or hidden in the settings menu.
• All the contents in the Extra Column in the printed version are incorpo-
rated into the entry.

- The definitions are displayed in blue, while every example is displayed per line in italicized black, which enables us to make a clear and quick distinction between them.
- The stress for compounds, which is shown as an underscore on the printed version, is omitted. This can be a serious shortcoming.
- Most of the front and back matters are included in the on-line help of the CD except for the pictorial illustrations and the supplement "Access to English" in the back matter. There are some materials available only in the CD, and some others are explained in more detail in the CD; the CD offers an article "Analyzing the Data" which gives an example of concordance lines and how the collocates can be obtained from the data. As for the lacking data, illustrations and "Access to English", it may be true that their absence does not devalue the CD seriously, but their inclusion would be of use to some users, at least doing no harm.

7.2.2. COBAm-CD

The dictionary contents are displayed basically in the same manner as the printed version, but there are some differences:

- IPA transcriptions are not included.
- The examples are indented and displayed per line in italicized red, which enables us to make a clear and quick distinction between definitions and examples.
- While the boxes headed as Picture Dictionary, Thesaurus, Usage, Word Link, Word Partnership, and Word Web are incorporated into the body of entries in the printed version, the CD-ROM version just offers their corresponding buttons in the entries which display the boxes when clicked. This reduces the accessibility to the information provided in the boxes. It is, in contrast, useful that the Word Web & Picture Dictionary tab provides us the complete list of the photographs and drawings.
- The sense numbers in the Thesaurus and Word Partnership boxes are omitted in the CD-ROM version. Figure 7.3 is a Word Partnership box under condition, where the sense numbers available in the printed version are not displayed. They show which sense the base word ("condition" in this example) is used in. It is not clear why the sense numbers are not given.

- Verbal patterns and grammatical information such as 
  
- For the homographic words that have more than one entry, the collections of the summary of the entries, called menus, are treated as headwords in addition to the entries themselves. For example, there are three entries for "lead"; the second one corresponds to lead, the third one is lead, and the first one includes both as well as the summaries (see Figure 7.4). The biggest problem is that only the first

![Figure 7.3 Word Partnership box in COBAm-CD](image)

![Figure 7.4 Homographs in COBAm-CD](image)
entry has the feature buttons such as Thesaurus. This is highly confusing to the users.

- The references to other entries, which are indicated as “→ see . . . ”, are omitted when they are placed after definitions. This is extremely undesirable. The references to other compound entries are not hyperlinked to the entries in question; examiner has “see also: medical examiner” and when we want to look it up we have to put the compound into the search box and press “Enter”. This is very inconvenient.

- When we are using the Dictionary mode, Word Webs and Picture Dictionaries are accessible only from the headwords that offer the boxes in their entries. Although the boxes have many related words, their entries have no link to the boxes. This is very dissatisfying.

- The explanatory words in the drawings and photographs cannot be double-clicked for searching them.

- Frequency information available in the printed version, such as ••••, are omitted.

- None of the front and back matters in the printed version, except for Activity Guide and Brief Grammar Reference, are included in the CD. Some users would find it useful if the CD provided just the same information as in the printed version.

7.3. Search

The searchability in the CD-ROM versions is much higher than that in the printed one where only search by entry items is possible. Some major characteristics of the search functions available in COB5-CD and COBAm-CD are described below.

7.3.1. COB5-CD

When one or more letters are typed in the search box, the nearest matching entry beginning with those letters is highlighted in the list of the headwords. By selecting an item from the headwords in the list, its definition is shown in the main entry section.

Typing a word or phrase in the search box and clicking the “OK” button or pressing the “Enter” key switches the left section to the Result tab where the following items are displayed:

- Compounds: All the compounds including the search key as a component word are listed under this heading. The morphologically related items with the key are also displayed under this heading; if “good” is searched for, the resulting list contains not only good day and good evening, but also brown goods, goodbye, and best-selling, some of which are impossible to find simultaneously in the printed version where all the entry items are alphabetically arranged.

- Phrasal Verbs: All the phrasal verbs (available in the headword list) including the search key as a component verb are listed under this heading. We cannot, however, search for the phrasal verbs with prepositions or adverbs.

- Full Text: This section displays the result of the search performed through the whole body of the dictionary, consisting of the following subsections:
  1. Definitions: All the definitions including the search key are searched for.
  2. Extras: Although the on-line help says this part “displays all usage notes and cross references which contain the word you have searched for”, it actually displays the notes on Americanisms and Briticisms, geographical labels such as BRIT and SCOTTISH, and style labels such as COMPUTING and HUMOROUS, but not other ordinary usage notes.
  3. Examples: The search key is searched for from all the dictionary examples and WordBank examples. The search result from the dictionary examples displays the headword as well as the example (see Figure 7.5); jumping to the example in question in the entry would enable us to know in which sense of the headword the search key is used in the example.
  4. Synonyms/Antonyms: The range of the search is all the words given as synonyms (given after = symbols in entries) and antonyms (given after ≠ symbols). For example, one of the synonym candidates of way is blaze, whose seventh definition “If someone blazes a trail, they discover or develop something
new.” has “lead the way”. It is likely that users will misunderstand that they can search for the synonyms and antonyms of the search key.

Full-text searches can be performed quickly.

Headwords with accented characters, such as à, can be searched for by altering accented characters with simple alphabetic characters. Hyphenated headwords can be searched for either with the complete form, with compositional parts with hyphens altered by spaces, or without hyphens nor spaces.

When the search key does not match any item, the nearest matching entry is highlighted in the headword list under the Index tab in the left section. Entering a spelling not included as a headword and clicking "OK" or pressing "Enter" displays the candidates under the Result tab.

If “Morphological search” option is ticked in the settings menu, searching with inflected forms yields the correct result for the lemma whose inflected forms are not included as a headword, although the headword is not highlighted in the list under the Index tab.

Inflected forms, speaker icons, and parts of speech are not displayed in the Quick View window. The information available in “menus”, where homographs or polysemous words are split into more than one entry and given brief explanations for each entry, is not displayed in the Quick View window, either; lead has a menu consisting of 1 BEING AHEAD OR TAKING SOMEONE SOMEWHERE and 2 SUBSTANCES in the printed version, but this information is omitted in the Quick View window while the usual Dictionary mode successfully shows it. Highlighting does not function, either.
When there are several headwords for the double-clicked word, the first candidate such as in and lead is displayed. The window cannot be resized, which lowers the usability. When a word is double-clicked within the Quick View window, the window shows the new entry. Double-clicking on an inflected form displays its lemma. Clicking the “Go to dictionary” button changes the application to the Dictionary mode where the entry is displayed in the main section. Although this Quick View function is essentially very useful, there are some minor problematic features. One example is that when one double-clicks on “means” used as a verb, the entry of means as a noun is searched for. Another example is that double-clicking on “case” in “= suit, case” seen in the third definition of action searches for the entry of suit.

Clicking on the v-shaped button next to the “OK” button displays the latest 32 items in the search history. The program also offers the two arrow buttons for going back and proceeding again, which is particularly useful when we want to look up an item that was searched for just a little time ago, for example. The history is shared between the dictionary and WordBank, which is also of use.

How the search works when the key is a phrase is not clear enough. It seems that the exact phrase or all its inflected versions with the component words in the same order are searched for through definitions, “Extras” (mostly notes on Americanisms and Briticisms), dictionary examples, WordBank examples, synonyms, and antonyms; each component word in the search key is searched for through compounds and phrasal verbs. Therefore, idiomatic phrases are difficult to search for; when one wants to know the meaning of the expression “wash one’s hands of . . . ”, but does not know that this is a fixed expression, and tries to search with the key “wash hand(s)”, then the correct result cannot be obtained.

The following types of searches are not possible in COB5-CD, which is quite unsatisfactory:

- Wildcard (* and ?) searches cannot be performed; only prefix-match searches, such as “nation-”, are available. The CD-ROM versions of other major EFL dictionaries such as LDOCE4 and OALD7 allow wildcard searches. Substring matching, such as “-ize-”, is not possible, either. Boolean searches using AND/OR/NOT cannot be performed. These limitations lower the searchability.

- It is not possible to search for a word, phrase, or string with wildcards within an entry. This is a serious drawback when the entry is large and the amount of text is vast.

- Searching with pragmatic labels such as “approval” cannot be performed.

7.3.2. COBAm-CD

7.3.2.1. Basic search

When one or more letters are typed in the search box, the candidates beginning with the search key string are shown in the list below the search box. By clicking on an item from the headwords in the candidates, its definition is shown in the main entry section. This electronic version does not include a spell-check function; when the search key does not match any headword, no candidate is shown. Since it may be troublesome for some users to be required to click on an item with a mouse, pressing the “Enter” key for starting the search should be possible as well.

Headwords with accented characters cannot be searched for by altering accented characters with simple alphabetic ones; we have to put the beginning part before the accented characters into the search box and select the appropriate headword. It is highly problematic when the word we want to look up starts with an accented character as in the case of à la; we have to enter any one character, delete it, and look for the headword in question while scrolling up and down the headword list from A, a to zygote. This is far from user-friendly.

Double-clicking on any word within the definitions and examples will execute a search in a small window, which will be called a “small lookup window” henceforth (see Figure 7.7). Inflected forms, speaker icons, and feature buttons are not displayed in the small lookup window. When there are several headwords for the double-clicked word, the first candidate is displayed as in the case of COB5-CD; however, when there are several homographic headwords available, we do not have the problem which
happens in COB5-CD that the information we want to know is not shown since the first entry includes all the homographic entries. It is also unfortunate for the users that the small lookup window cannot be resized. When a word is double-clicked within the small lookup window, the window shows the new entry. Double-clicking on an inflected form displays its lemma. Clicking the “Show Definition” button changes the application to the Dictionary mode where the entry is displayed in the main section. Double-clicking on a hyphenated word searches for a clicked portion split by the hyphen(s), not the whole word.

Clicking “Show History” displays the latest 25 items in the search history. The program also offers the two arrow buttons for going back and proceeding again just as COBS-CD does, which is useful.

Highlighting the search key is not provided, nor is it possible to search for a word, phrase, or string with wildcards within an entry, which can be troublesome when we have to read through long entries.

**7.3.2.2. Advanced search**

In the advanced search mode, we can perform searches with more detailed options (see Figure 7.8).

For headwords, prefix-match, suffix-match, substring-match, and ex-
user-friendly when the entries are long since the program does not highlight the search key and does not provide a within-entry search function.

Although wildcard (\* and ?) searches cannot be performed, most searching needs will be satisfied through prefix-match, substring-match, and suffix-match searches. Boolean searches using AND/OR/NOT are not possible, and some idioms are difficult to find just as in the case of COB5-CD described above.

To sum up, the occasions where the above functions are utilized to the users' convenience are limited.

7.4. WordBank

This is one of the most attractive features of COB5-CD. It would be more useful if WordBank was available in COBAm-CD, too, since COBAm also uses the data from the Bank of English, on which WordBank is based.

Searching for a word or phrase displays the first 200 examples in the main entry section, while the left section displays the number of examples available divided into groups of 200, clicking on which makes the entry section show the corresponding examples (see Figure 7.9).

![Figure 7.9 WordBank in COB5-CD](image)

It is regrettable that only the simple word/phrase search cannot be performed, although the program automatically searches for all the inflected forms; availability of wildcards and AND/OR/NOT would enhance the searchability. Each example is labeled US/UK and written/spoken; the broadest differentiation, such as journalism or fiction, would be desirable so we could make more effective use of the data. The dictionary and WordBank are fairly well integrated with each other.

7.5. Recorded pronunciations

Both COB5-CD and COBAm-CD provide pronunciations for all headwords including compounds. Inflected forms are also recorded.

7.5.1. COB5-CD

Clicking on the speaker icon plays back the recorded pronunciation. If "Pronounce every entry" option is ticked in the settings menu, the first of the recorded sounds is automatically played as soon as the entry is displayed.

Both British and American pronunciations are available, with blue speaker icons for British pronunciations and red icons for American ones. For words that do not have variation between the two varieties, only the British pronunciation is provided. For example, there is only one recording for heart, and it is not rhotic (see 3.2.1 for rhoticity). COB5 can thus be said to favor the British variety. This could be misleading for users who are not familiar with English accents because they could misunderstand that both the American and the British pronounce the word in the same way as the recording.

When there is a note on pronunciation, the pronunciations within the note are supplied with recorded sounds, but not the headword itself. This causes incompleteness in some examples. For example, there is a note for conduct.
indicating different pronunciations depending on whether it is used as a verb or noun, but no recording is available for the inflected form “conducts” (see Figure 7.10).

COB5-CD also offers a function where the user can record his or her own pronunciation and compare it to the recorded pronunciation that was most recently played back.

7.5.2. COBAm-CD
Clicking on a speaker icon or inflected form gives the recorded pronunciation. However, IPA transcriptions are not provided in the CD-ROM version. Thus, it may be difficult for users to know which syllable carries the stress. This is a serious problem because users have to listen to the recording in order to verify the pronunciation.10)

When there are more than one variant (e.g. two for room and three for eczema), these variants are all pronounced. However, it is not made apparent that there is more than one variant until we listen to it since there is no IPA transcription provided.

Pronunciation notes are omitted; although the printed version has the same pronunciation note for conduct as available in COB5-CD, it is not available in COBAm-CD. This deprives users of the opportunity to learn the different stress patterns depending on parts of speech.

7.6. Overall evaluation of the CD-ROMs
As for COB5-CD, the interface has been refined from the previous version, COB4-CD, in most aspects, and the overall impression is fairly good. In a practical aspect, it is convenient that search functions are available without installation, and the CD-ROM is not needed after we have installed it on a hard disk. One serious shortcoming is, however, that the search function is poor; unavailability of wildcards, poor searchability for phrases, inability to perform label searches, should, in particular, be improved.

COBAm-CD is, in contrast, just an electronic viewer of COBAm with some searchability normally associated with computers. Although installation on a hard disk is unnecessary just like COB5-CD, the CD-ROM is always needed to run the program, which is inconvenient. It is also disappointing that even basic functions are not offered; search keys are not highlighted in the entry, and we cannot copy nor print out a whole entry. Some users would hope to use WordBank in COBAm-CD as in COB5-CD. As for the interface, COBAm-CD needs much more innovation. It is desirable that the interface and search functions work just as the users expect them to work; the gap is large between what is available in the current software and our expectations. In sum, there is little to be praised in COBAm-CD compared with other CD-ROM dictionaries. We have to say that this CD-ROM version is far from satisfactory.

If we use the two CD-ROM dictionaries at the same time, we very soon find that their interfaces and functions greatly differ from each other, and there is no link between them, which is unfortunate for the users.

In general, even if the contents are the same in the printed and CD-ROM versions, there is yet a large quantity of information retrievable only in the electronic version; it is possible only in the CD-ROM versions to perform flexible searches through definitions and examples, and to search for all the headwords or definitions with certain labels. The greatest advantage of the CD-ROM version is to be able to provide such searchability; both COB5-CD and COBAm-CD have room for improvement to achieve this goal.

(Section 7 by Ishii)

8. Conclusion
Although COBAm has much in common with the Collins COBUILD dictionaries, the overall structure of the dictionary has slightly changed. The following are the result of our collaborative study of COBAm.

Section 2 discussed the differences between COBAm and COB5 concerning headwords, run-ons, phrases and phrasal verbs, superheadwords and Frequency Bands. As for the headwords and run-ons, a sampling of about 10% of the dictionary body has revealed that a considerable number of American terms have been newly introduced in COBAm, whereas many British items have been deleted. Concerning the headwords, some have been changed to American spellings. For example labour in COB5 appears as labor in COBAm. In addition, some changes have been
observed with respect to the use of hyphens and periods. Regarding the phrases and phrasal verbs, a noticeable number of items have been deleted in COBAm while few have been added. A problem that has been pointed out concerning this rather drastic deletion is that users may be unable to find explanations for the phrases or phrasal verbs used in the definitions and examples of COBAm. In addition, as a result of the deletion of the Extra Column, it is difficult to find phrases and phrasal verbs in COBAm. Superheadwords in COBAm are much the same as in COB5, but the shortcuts are sometimes user-unfriendly and the application of superheadwords is inconsistent. With regard to Frequency Bands, only minor changes have been made except that they have been assigned to calendrical and geographical terms in COBAm.

In Section 3, COBAm's pronunciation was reviewed, focusing on whether it accurately represents American English. It began with a general discussion (3.1), in which it was noted that the dialect(s) of description may not be straightforwardly understood, since it appears to incorporate RP. Variation within GenAm is not discussed in the Guide (unlike LAAD2); variation may be recorded in the entries, but these variations can be interpreted to be RP forms, rather than other variants in GenAm. Allophonic features were discussed (3.1.3); COBAm, like COB5, presents very few allophonic features, while LAAD2 gives more, especially when they are characteristics of American English (e.g. tap, t/d deletion).

COBAm's American-ness was then examined with respect to selected diagnostic features, first in the paper version (3.2). The features can be divided into three groups, varying in the degree to which they are in accordance with the literature. To start with the non-accordance end, the merger of /o/ and /a/ is generally assumed in the literature, while COBAm maintains the contrast. There are two features that are not expressed in COBAm but are described as change in progress in the literature: merger of /a, o, 3/ and neutralization between /æ/ and /e/ before intervocalic /r/. The remainder mostly appears as expected from the literature (e.g. rhoticity, loss of /h/ in the onset /hw/ cluster), non-contrast between two as before tautosyllabic /r/, neutralization between /æt/ and /æ/ before intervocalic /r/), though on rhoticity, the transcription is systematically bilectal, which was pointed out to be confusing to dictionary users, who expect American forms only. Some of these features (ask words, yod dropping, non-contrast between /ə/ and /æ/ before /t/, lexical incidences) further show variation in certain words, though the non-GenAm forms may not be intended to represent GenAm in COBAm, which, again, is confusing to the users.

On the CD-ROM (3.3), a subset of the features was diagnosed. The use of cluster reduction in /hωw/ is fully consistent with the paper version of COBAm and also expected from the literature. Yod dropping is consistent with the paper version except in one word. Ask words show a larger discrepancy from the paper version, with more instances of alternative vowel /a/ (in addition to the expected /æ/) than in the paper version. Certain contrasts that are described in the literature as disappearing in GenAm but maintained in the paper version were then discussed (merger of low back vowels /o, a, 3/ and neutralization between /æ/ and /e/ before intervocalic /r/). Impressionistic analysis of certain (near) minimal pairs found possible absence of these contrasts. Acoustic analysis was conducted on /ɔ/ vs. /o/, and /æ/ vs. /e/ before /r/. Between /ɔ/ and /o/, a consistent but (very) small differences in F1 (height) were found. In /æ/ and /e/ before /r/, these two vowels were found to be fairly close to each other compared to /æ/ before /d/, especially in F1, suggesting acoustic neutralization of /æ/ and /e/ before /r/, especially in vowel height.

In Section 4, it was pointed out that COBAm does not explicitly state how it actually uses the DV, but that its explanation suggests that it uses words indiscriminately in its definitions, even those from outside of the DV list. In fact the dictionary uses words and phrases not listed as part of the DV. We can safely say that COBAm does not use a DV in its strict sense. COBAm's claim to use a DV may be at least to some extent motivated by commercialism because DV has become an important selling point in the highly competitive market.

Concerning sense descriptions, the vast majority of definitions in COB5 and COBAm are almost the same except for the following changes: (1) systematic replacements of Britishisms with the corresponding Americanisms; (2) attempts at more readable and/or more informative explanations;
(3) systematic reflections of political correctness. Although these changes are welcome, it is regrettable that COBAm has a far narrower coverage of senses and that the dictionary's policy on its selection and arrangement of senses sometimes seems arbitrary.

With respect to labels, again there are no fundamental differences between COB5 and COBAm. Although COBAm's attempt to simplify the geographical labels may be a step forward, there seems to be some confusion between the style labels and the pragmatic labels. It should also be mentioned that the abolishment of the Extra Columns has made the pragmatic labels less conspicuous. There is some room for improvement in COBAm's use of labels.

In Section 5, through the comparison between COBAm and COB5, we have found that a large number of examples have been deleted in COBAm while only a small number of examples have been added. Many examples have been changed simply from British place names to American place names, Briticism to Americanism, etc. There are some examples whose changes can be regarded as improvements when they are replaced by new examples. The majority of COB5's examples, however, are used in COBAm without any change.

COBAm was also compared with LDOCE4 and OALD7. This comparison has clarified that COBAm offers far fewer examples than the other two dictionaries. Many more examples should be added to represent the wide range of expressions in actual use. COBAm's examples tend to be more difficult to understand because they are written using complex syntactic structures and difficult phraseology. The topics of the examples are often unfamiliar to learners. They would be suitable for a general purpose dictionary, but in a pedagogical dictionary, they should be changed to more familiar ones in easy phrasing.

In Section 6, the features of Vocabulary Builders (VDs) were described. VDs consist of six boxes: Picture Dictionary boxes, Word Webs, Word Partnerships, Word Links, Thesauruses, and Usage.

Unlike the illustration pages that are usually bound together in the appendices or in a middle [back] matter in traditional learners' dictionaries, Picture Dictionaries (PDs) are inserted in the body of the dictionary no matter how large and space-consuming they are, which makes them easy to access for dictionary users. The words described in PDs are mostly nouns. They are not necessarily the academic words that are emphasized in the front matter. Many of them are words that benefit from the provision of encyclopedic information.

Word Webs (WWs) probably are the most innovative aspect in the VDs. They are intended to give a deeper understanding of a word or a concept by providing related encyclopedic information along with the key words that are thematically related to the headword. There are cross-references to WWs, which also help the users to better understand the meaning and usage of the headwords.

There are boxes that are similar to the Word Partnership boxes, Word Link boxes, Thesaurus boxes, and Usage boxes in other learners’ dictionaries. Therefore, they are not original innovations in COBAm. However, they may look more eye-catching in COBAm than in other learner’s dictionaries, since they are colorful and stand out on pages that are basically black and white.

Our comparison and analysis of the differences between the printed and the CD-ROM versions of COB5 and COBAm in Section 7 have revealed that COB5-CD is user-friendly enough in many respects while COBAm-CD leaves much room for improvement to be a good electronic dictionary for many users.

NOTES

Section 2
1) The estimated numbers are calculated based on the number of the pages in each edition. The total number of pages of the sampling in COBAm is 165 (30 + 82(A) + 53(L)), whereas in COB5 it is 178 (30 + 85(A) + 63(L)). As for the preliminary survey, the number of corresponding pages in COB5 is not exactly 30 pages, but no adjustment was made because the number of pages is roughly the same amount. The average per page is then multiplied by the total number of pages in the dictionaries, 1,522 and 1,695 for COBAm and COB5 respectively.
2) Some of the headwords such as labour, labour camp and labour force become empty entries while others are simply replaced by the American spellings (e.g. lackluster and left-of-center).
3) These types of changes are also observed in OALD7. See Komuro et al. (2006: 62-64) for detail.
form in all three dictionaries segmentally, but the primary stress falls on the penultimate syllable, as opposed to what Wells gives (first syllable). *LAAD2* gives two forms in *borough* and *thorough*—/ɔ/ and /ə/ in the first vowel, still different from Wells’s RP form in the last vowels. This variation in the first vowel seems to be the same kind as discussed in 3.2.1.3, i.e. /ɔ/ and /ə/ before /t/. *antti* has a form with /i/ at the end in all three dictionaries (along with GenAm form in *COBAm*, as the only form in *COB5*, and along with GenAm form and RP form in *LAAD2*). In *COBAm and COB5*, this might be interpreted as corresponding to Wells’s RP form /-ti/ (as opposed to the GenAm form /-tai/) since in Wells the unattested vowel at the end of a word such as *happy* is assigned to /i/ in RP, which can be translated to /i/ in these dictionaries since the last vowel in *happy* is /i/, not /ɪ/. However, in *LAAD2*, this interpretation does not hold since all three possibilities are shown, i.e. /ɛnti/, /ɛntaɪ/, /ɛnti/. 18) DARE includes /-dim/ with the tag “also” (*cf. /-,din/ with “usually”). However, it is not known from DARE whether this form is used in current GenAm. It is reported to be heard in New England, especially in Rhode Island, with the source being *Linguistic Atlas of New England* (1943) and Kurath and McDavid (1961), but it is not listed in other sources, including the recordings that DARE made in 1965–70, which covered a large area in the States. (It is also not known if the form was still used in Rhode Island at the time of DARE interviews, since RI is not listed for any variant from the DARE data.) 19) See Komuro et al. (2006: 72–73) for possible merger in the recording of OALD7. 20) The spectrograms show some periods where information is missing from certain frequency ranges. However, the frequency range for the first two formants is not affected in this way, so we proceeded in the acoustic analysis. 21) Hairy and caring are transcribed as /ɛə/. We assume this to mean /ɛə/ in GenAm as discussed in 3.2.1. 22) This happens to give comparable duration also. Approximately 106 ms was taken in hairy, harry and merry, aprox. 137 ms in marry, approx. 81 ms in caring, and approx. 87 ms in marry. Cf. The duration for the other vowels is 100 ms.

Section 4

1) As long as OALD7, another EFL dictionary of almost the same size, claims to use 3,083 items as part of its DV (The Oxford 3000 and Language study terms (p. R99 and p. R113)), COBAm’s DV cannot be too large. COBAm therefore should have provided the user with sufficient information so that the user may judge the dictionary.

2) Both COB5 and COBAm put a note to the use of “they” under the entry for the pronoun: “Some people think this use is incorrect”, which may show how difficult it is to keep a balance between descriptiveness and prescriptiveness in the EFL context.

3) Although in the front matter COBAm lists the seven pragmatic labels in lower case with the square frame following its predecessors like /approval/, within entries it uses small capital letters for the labels with brackets like [APPROVAL]. The list of the labels in COBAm might have simply been taken from its predecessors.

Section 5

1) The changes have by no means been complete. In the example above, “autumn” in: In autumn 1944 the Allies invaded the Italian . . . , remains unchanged while “allies” has been
capitulated. Similarly, "railway" in The railway strike fizzled out on its second day as drivers returned to work, and "film" in: Not many films are made in Finland, have not been replaced with "road" and "movie" respectively.

2) There are some mistakes. The example of the phrase “as many as” in COB5 is As many as four and a half million people watched today’s parade. This example has been changed to 4 million people watched today’s parade in COBAm dropping the very phrase the example is meant to illustrate. In the CD version, it has been corrected to New York City police say that as many as 4 million people watched today's parade.

3) LDOCE4 says it "include 40 per cent more examples in this new edition of the book" (p. x).

4) This example does not belong here. It may have been misplaced.

5) Similarly, “it” in It could be a bit problematic, legally speaking.

6) The total number of examples in each of these entries is as follows. (Colloquial examples in OALD7 are included.)

<table>
<thead>
<tr>
<th></th>
<th>COBAm</th>
<th>LDOCE4</th>
<th>OALD7</th>
</tr>
</thead>
<tbody>
<tr>
<td>door</td>
<td>14</td>
<td>19</td>
<td>30</td>
</tr>
<tr>
<td>money</td>
<td>11</td>
<td>30</td>
<td>35</td>
</tr>
<tr>
<td>eat</td>
<td>10</td>
<td>33</td>
<td>22</td>
</tr>
<tr>
<td>see</td>
<td>40</td>
<td>144</td>
<td>81</td>
</tr>
<tr>
<td>whether</td>
<td>4</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

7) It gives examples such as ‘The Indians are our subjects’ from the 1911 edition of COD, ‘Give her a good beating’, and ‘A drunken woman is a deplorable spectacle’ cited from Allen (1986: 2).

8) LDOCE4 says: Our editors and lexicographers edit them (= examples) to remove difficult words and to ensure they are comprehensible examples of that meaning of the word within a single sentence or phrase (p. x).

9) Four words: “principal”, “brochure”, “superb”, and “overhead” are not on the list.

10) They have been adapted to fit in the dictionary: “The majority of the examples in the dictionary are taken word for word from one of the texts in The Bank of English. Occasionally, we have made very minor changes to them, so that they are more successful as dictionary examples” (COB5: xv).

Section 7

1) The entries registered by a user can be searched for only through headwords; the searches for definitions and examples available in the advanced search do not work for the user-added entries.

2) There are some cases where the entries including the notes or illustrations are different between the printed and electronic versions, and some pieces of information are available only in the CD-ROM version (see 6.1 and 6.2).

3) This function has not worked properly in the present writers’ environment.

4) Double-clicking on the compound can jump to either medical or examiner.

5) This highlighting function does not work correctly for compounds with spaces between the component words.

6) Searching for compounds with apostrophes by typing in the component words including apostrophes does not yield the correct results, nor can the words in question without apostrophes. For instance, Adam’s apple is not given under the Compounds section when the search key is “Adam(s)”.

7) The label TRADEMARK is an exception; when it is searched for, the entries including the label are shown under the subsection of Definitions, the reason of which is not known.

8) The is not highlighted in the entry of the, perhaps because there are too many candidates.

9) Clicking on “means” used as a verb searches for means as a noun in the same way as COBS-CD.

10) Those compounds that are without IPA transcriptions but underscored for stress markings in the printed version are shown in the same manner in the CD-ROM version.

Dictionaries


REFERENCES


ANAE = Labov et al. 2006.


