

An Empirical Approach to the Review of EFL Dictionaries: *LDOCE2*, *LDOCE3*, *COBUILD1*, and *COBUILD2**

Shigeru Yamada

1 Introduction

This paper attempts to critically and comparatively evaluate EFL dictionaries. The approach is empirical: a test was devised to look into particular parts of the dictionary consultation process for comprehension and the test was given in class with college students as subjects. As lexicographical resources, the photocopied entries of *LDOCE2*, *LDOCE3*, *COBUILD1*, and *COBUILD2* (see Appendices¹) were used². The purposes, methodology, advantages and limitations, and findings of the study follow.

2 Purposes

The main research purposes of the study are to observe the following aspects of dictionary consultation (cf. Scholfield 1982, 1999; Hartmann 1989, 2001):

- (1) how users get to the appropriate sub-entry
- (2) how they interpret the definition
- (3) how they use the elicited information to understand (translate) the target lexical item in the original sentence.

3 Methodology

3.1 Pilot study

A pilot study was conducted to check the feasibility of the study and finalise the research design. Two B4-size sheets (Appendix 1) were prepared. Both include common directions and questions on the right-hand side with two photocopied entries of different dictionaries on the opposite side. One had those of *LDOCE3* (in the top left-hand corner) and *LDOCE2* (in the bottom right-hand corner), and the other those of *COBUILD2* and *COBUILD1* arranged in the same way. Two classes (65 students in total) participated in the study. They were divided into four groups, so that each used a different dictionary. Class A was given the *LDOCE* sheet, and 17 students consulted the entry of *LDOCE3* and the rest (16) that of *LDOCE2*. Class B was provided with the *COBUILD* sheet, and a half (16) used the entry of *COBUILD2* and the other half that of *COBUILD1*.

Subjects were female college freshmen, reading English. They had had six years of formal English language education. There were three students in Class B who had lived abroad for more than one year³. All were experienced in using learners' dictionaries between English and Japanese (especially English-Japanese ones). There were three students who had used monolingual English dictionaries⁴.

The pilot study mainly required the students to look up a designated item in a given sentence in the attached photocopy of an EFL dictionary entry in an effort to (better) translate the sentence. They were given two chances of translation before and after consulting the dictionary entry. The study was conducted at the beginning of each class by the following procedure. The test sheet was handed out face down. The students had to turn it over and fold it in half, so that they could only look at the directions side when they translated the sentence for the first time. They did this without knowing which word in the sentence was the target item to look up later. In the pilot study the sentence to be translated was:

[Student to professor]

I'd like to talk to you about the assignment you gave us today.

The target word was *assignment*. In what follows, the procedure is explained, task by task, as each appears on the sheet.

Task 1: Translate the following English sentence into Japanese.

This initial translation task was intended to check the students' prior knowledge of the target word. This is important to note because in real life, if they already know the meaning of a word, they would not go to a dictionary (for the meaning at least). Moreover, familiarity and unfamiliarity with an item may influence the reference—speed, strategies to be adopted, and the understanding of the definition. In the analysis, cases in which students already knew the meaning of the target word have to be separately considered from those in which they did not.

Task 2: Consult the dictionary extract on the left-hand side for the meaning of the target word from the above sentence. Indicate a corresponding sense number from the dictionary entry in the space provided. Press the PLAY button when you start your consultation and press the STOP button when you have finished. Enter the figures on the counter in the space provided to indicate your reference time.

The students worked through Tasks 2-5 on their own. From this point onwards, they were allowed to look at the dictionary entry on the other side. By looking at the attached entries on *assignment*, they first recognised it as the target word. In this task the students looked up the meaning of the target word in the dictionary entry, timing how long their reference took. Since both classes conveniently took place in the language laboratory,

the tape counter was used for the latter purpose⁵. As long as the dictionary is a tool for quick reference, a good dictionary should serve this purpose. Reference time can be used as a yardstick to determine the difficulty of reference and the quality of an entry. It may be interesting to observe the students' reference with respect to speed over a period of time. They provided the result of their reference by stating the sense number they had chosen as appropriate. This was used as a basis for judging whether or not their consultation was successful.

***If you realise in the process that you have given the wrong sense number, cross it out, enter a second number in the square brackets, and continue with your work.**

Dictionary consultation can be such a complex process and does not necessarily go without a hitch. This can occur even more frequently at an early stage of using an L2 monolingual dictionary. It might happen that students changed their minds after deciding on a meaning and entering the sense number. This asterisked part was specially provided to cater for such likely, not smooth references⁶. In the analysis, cases like this should be treated separately from those where reference went smoothly at one go.

Task 3: How did you decide on the sense number above? State your reference process in this case on a step-by-step basis.

This task was to probe into the students' actual consultation procedure. They were asked to give a detailed account of how they got to the meaning of their choice, tracing it from short-term memory. In order for the researcher to elicit relevant information, each group was furnished beforehand with general knowledge about the microstructure of each of the dictionaries involved.

Task 4: Translate the English sentence above (under Task 1) into Japanese again. If it is exactly the same as your first translation, just state so.

Here the students were asked to translate the same sentence a second time after consulting the entry. The aim was to find out what impact the dictionary consultation had on their understanding of the meaning of the target word as reflected in their second translation.

Task 5: Translate into Japanese the definition of the sense number you chose in Task 2. You do not have to translate any items you do not understand (including codes and abbreviations) but underline them in the dictionary extract.

This task was intended to assess the students' understanding of the definition and pinpoint problem areas—the words and constructions they have difficulty in understanding. The use of translation for this purpose may not be free from criticism: translation involves two activities (i.e. understanding the source language and putting it into the target language) and therefore it is hard to detect at which stage an error occurs if it does. However, there seem to be few problems for Japanese college students, who are trained at high school to understand English through translation into Japanese, so much so that understanding English is somewhat synonymous with translating into Japanese. In order to respond to the possible criticism, however, measures have been worked out. Task 6 is set to check the first phase of translation (comprehension of the definition). The students' translation in Task 5 is going to be used mainly to identify complexities in the definition.

Before Task 6 I let the students know the appropriate sense number for each dictionary: it was Sense 1 for all, except *COBUILD1*, for which Sense 1.2 had to be picked out. Then I gave an example translation of the English sentence.

Task 6: Check your understanding of the definition against the instructor's translation. Evaluate your understanding by circling one:
a Perfect understanding b Half understanding c No understanding
d Between a and b d Between b and c f Other

The students checked their understanding of the relevant definition against the example translation I gave orally for each dictionary. And they rated their comprehension by circling one out of **a-f**. The students who made the wrong choice of a definition were instructed to circle **f**. This way, the students' evaluation here is guaranteed to reflect the first phase of the translation process only—understanding of the definition.

Lastly the students were asked to make comments on the entry they used (in comparison to the other entry of the same dictionary) and on their consultation. I posed a few questions to draw the students' attention to particular points.

3.2 Final research design

The research design for the study was then finalised, based on the results of the pilot study. One addition and one substitution were made.

Task 3.1: Why did you choose the above sense as appropriate? State the grounds for your decision.

Task 3.1 asked about the factors which led students to choose one meaning over the other(s) from the entry. This was added while Task 3 in the pilot study remained as 3.2 in the revised design. These two problems are so closely related that some students touched upon what had prompted them to choose that sense in Task 3 in the pilot study. Since the justification of their decision on meaning can constitute such crucial feedback (especially in consulting an entry with only a few senses, where (much of) strategic

reference is not required) it needs to be securely elicited.

Task 6: Check your translation of the definition against the instructor's translation. Evaluate the accuracy of your translation by circling one: a 100% b 99-90% c 89-64% d 63-37% e 36-11% f 10-1% g 0%

The multiple choices for Task 6 were "refined" with percentage points. However, this is not a substantial change from the pilot study. It is impossible to gauge one's understanding of a definition in exact numerical terms, anyway. The students understood the choices as follows: **a** represented 'perfect understanding' and **b** 'almost perfect understanding', **g** indicated 'no understanding' and **f** 'almost no understanding'. The rest was divided into three with **d** pointing to the neighbourhood of 50% understanding⁷.

The final version of the test sheet is found in Appendix 2. Given as a part of an English class (90 minutes), the whole process of the test had to be kept under 20 minutes. It was thus decided that one and a half minutes were allotted to Task 1 and seven minutes to Tasks 2-5.

4 Advantages and limitations

This study is a simulation, focusing on particular parts of the dictionary consultation process for receptive purposes. Designed to be well-focused and not to overload subjects, it is not free from criticism for lack of authenticity. Advantages and limitations are both sides of a coin. The test imposed on the students the use of an EFL dictionary (entry) assigned by the researcher without allowing them the liberty of using a different-genre dictionary or even no dictionary use. Photocopied entries were used, rather than actual dictionaries, which inevitably makes this study microstructure-based. However, this made it possible to put four different versions of dictionaries into focus. The proportionate decline in the number of subjects per dictionary is a minus point but, on the other hand, it made a close examination from various angles possible. Thanks to the opportune availability of the

tape counter, reference was timed regularly and precisely, which is a key element of dictionary consultation (and its investigation and assessment), without worries that subjects might forget their watch or might not have one with a second hand or a stopwatch function. However, the time obtained in the test does not include that required for opening a dictionary and flipping through the pages for the appropriate entry, which takes place at the initial stage of the actual look-up. Since an English sentence which could be translated into Japanese in a short time had to be used, it may not have provided sufficient context, the kind available in real-life reading and conducive to good dictionary consultation. The students were allowed to look up only a designated vocabulary item in an English sentence in the test, which was enforced through the use of photocopied dictionary entries. Nonetheless, the target item may not have been the one they had to check and there may have been other words and phrases they wanted to consult. As the test was given as an in-class activity, the time allotted was inevitably limited; there is no guarantee that all students worked through the tasks without feeling a shortage of time. The research design depends on the students' translation of the English sentence in two ways—it is intended to check their prior familiarity with a target word and to estimate the impact of dictionary consultation as reflected in translation output. Nevertheless, there were a few occasions when translation did not turn out to be so useful for these purposes—it was not always easy to recognise a match between the English target word and its Japanese equivalent because of their discrepancies in parts-of-speech and nuances. Based on only the first of a series of thirteen tests, the study is detailed but of limited scope⁸.

5 Findings

5.1 The English sentence, the target item, and the appropriate sense number

The sentence to be translated in Tasks 1 and 4 was this:

I often feel nauseous on the bus.

The item to be looked up in Task 2 was *nauseous*. The appropriate sense number was 1 in the adjectival entries of *LDOCE2*, *LDOCE3*, and *COBUILD1*, where two senses are presented. *COBUILD2* provides only one sense, which fits the context.

5.2 Observational settings

Both editions of *LDOCE* define *nauseous* with reference to *nausea*. If subjects can not figure out the meaning of *nausea* in reading the definition of *nauseous*, they have to go to the nominal entry, which is a few entries above (or rely on the example[s]). To make this possible, the sheet for the *LDOCE* group included on the left-hand side the entries from *nausea* to *nauseous* from each dictionary (see Appendix 2.1). At the end of the study, I asked the students for their reaction to this treatment of *nauseous*.

COBUILD adopts different approaches to *nauseous*: the first edition provides two senses, while the second edition only one—that of the target word, probably owing to frequency discrepancies of the corpora utilised. Subjects were asked to state their preferences on this in their feedback.

5.3 Statistics on students' dictionary consultation and post-reference translation

Fifty-six students participated in this test: the number of subjects for each dictionary is as the furthest left-hand column of Table 1 indicates. The table summarises the statistics on the students' dictionary consultation (reference success/failure, time, and understanding of the definition concerned) and the rate of the appropriate post-reference translation of the target word. Those who were successful at sub-entry search are separately treated from those unsuccessful.

Table 1 Statistics of students' dictionary consultation and translation thereafter

Dictionaries (no.of subjects)	Reference success/failure	Reference breakdown	Reference time (sec.): Mean (quickest/slowest)	Understanding of definition	Right 2nd translation
<i>LDOCE3</i> (15 subjects)	Suc. 13 (86.7%) Fail. 2 (13.3%)	12 (80.0%) Rev ¹ 1 (6.7%) 2 (13.3%)	63.7 (26.8/94.4) 113.3 (46.6/180.0)	73.8% 12.5%	11 (84.6%) 0
<i>LDOCE2</i> (14)	Suc. 7 (50.0%) Fail. 7 (50.0%)	6 (42.9%) Rev ¹ (7.1%) 7 (50.0%)	41.2 (11.1/116.5) 90.0 (35.0/184.7)	56.4% 54.2% ⁴	5 (71.4%) 3 (42.9%)
<i>COBUILD2</i> (12)	N.A. ²	N.A. ²	N.A. ³	53.3%	10 (83.3%)
<i>COBUILD1</i> (15)	Suc. 11 (73.3%) Fail. 4 (26.7%)	10 (66.7%) Rev ¹ (6.7%) 4 (26.7%)	27.9 (8.2/65.2) 36.3 (22.1/64.1)	84.1% 20.0%	11 (100%) 0

1 "Rev" refers to those who "revised" their decision on the choice of the appropriate sense.

2 The referential success/failure distinction does not apply to the use of *COBUILD2*, which presents only the appropriate sense.

3 Neither does reference time.

4 One subject did not indicate the rate of her definition understanding.

I will explain what this table exhibits from left to right, taking the results of the *LDOCE3* group as an example, mainly. Among the 15 students who consulted the entry from the dictionary, 12 (80.0%) were successful, choosing the appropriate sense in Task 2, but two (13.3%) were unsuccessful. One (6.7%) made the wrong choice but later "revised" it. The success/failure distinction does not apply to the look-up of *nauseous* in *COBUILD2*, which gives only the appropriate sense.

As for reference time, those who changed their mind after finishing their look-up (those labeled "Rev" in Table 1) are ignored. The users of *COBUILD2* are not considered, either, because the entry was monosemous⁹. The mean reference time of the 12 subjects who succeeded in the look-up of *LDOCE3* was 63.7 seconds. Among these the quickest took 26.8 seconds and the slowest 94.4 seconds as presented in brackets following the mean. The average of the two students who failed in their sub-entry search was 113.3 seconds; the quicker was 46.6 seconds and the slower 180.0 seconds.

Based on the answers to Task 6, averages in the subjects' understanding of the definition of the appropriate sense could be worked out. In order to obtain a general picture, the calculation was made in the following way:

the understanding of those who circled **a** was determined to be 100%, **b** 95%, **c** 75%, **d** 50%, **e** 25 %, **f** 5%, and **g** 0%. Calculated on this basis, the mean understanding of the subjects who were successful in the look-up with *LDOCE3* was 73.8% while that of those who failed was 12.5%¹⁰.

Lastly, the relationship between reference success and post-reference translation is provided. Out of the 13 successful users of *LDOCE3*, 11 (84.6%) translated the target word appropriately in the post-reference translation (Task 4). On the other hand, both of those unsuccessful at sub-entry search failed to give any appropriate translation of the target word in the same task.

Table 1 provides a general idea of how the students performed in their reference and translation during the test I gave them. The overall observation is that, understandably, those who consulted the entry successfully were able to do so quickly, understand the definition well, and make proper use of the elicited information in translating the target word. Relevant parts of the data will be analysed closely and from different angles in the following sections.

5.4 Look-up strategies adopted

The consulted entries on *nauseous* with only one or two senses do not require sophisticated look-up strategies. Although there are a few students who ran systematic consultation methods even on such a simple entry, this section focuses on the results of Task 3.1. Table 2 shows the factors contributing to their final decision on meaning, and the number of students who mentioned them according to the editions, the dictionary groups, and the total. The breakdown of those successful and unsuccessful at reference (sub-entry search) is also indicated (in this order in brackets). In looking at the results of the *COBUILD* group (*COBUILDs*), it should be noted that the fact that all references of *COBUILD2* (with the appropriate sense only) are counted as successful has jacked up the rate of reference success (85.2%) (compared to that of *LDOCEs* [70.0%]).

Table 2 Factors contributing to students' final decision on meaning

	<i>LDOCE3</i>	<i>LDOCE2</i>	<i>COBUILD2</i>	<i>COBUILD1</i>	<i>LDOCEs</i>	<i>COBUILDs</i>	Total
	No. (ref. success/failure)	"	No.*	No. (ref. success/failure)	"	"	"
Definition	3 (2/1)	3 (1/2)	2	10 (6/4)	6 (3/3)	12 (8/4)	18 (11/7)
Example	6 (6/0)	6 (3/3)	2	1 (1/0)	12 (9/3)	3 (3/0)	15 (12/3)
Original sentence	3 (3/0)	1 (0/1)		1 (1/0)	4 (3/1)	1 (1/0)	5 (4/1)
Definition/example	1 (1/0)			1 (1/0)	1 (1/0)	1 (1/0)	2 (2/0)
Definition/example/original sentence				1 (1/0)		1 (1/0)	1 (1/0)
Other		1 (1 [label]/0)			1 (1/0)		1 (1/0)
No answer	2 (1/1)	3 (2/1)	8	1 (1/0)	5 (3/2)	9 (9/0)	14 (12/4)
Total	15 (13/2)	14 (7/7)	12	15 (11/4)	29 (20/9)	27 (23/4)	56 (43/13)

* Referential success or failure is not applicable because *COBUILD2* offers the context-fit sense only.

Overall, as the furthest right-hand column above shows, the greatest number of students (18) based their ultimate semantic decision on the definition, followed by the example (15 students). In terms of reference success, however, the example-oriented consultation outscored the definition-oriented 80% to 61.1%. It is of interest to observe that the students' preferences were sharply contrasted by consulted dictionaries: the greatest number of *LDOCE* users (41.4%) relied on the example, while the greatest number of *COBUILD* users (44.4%) turned to the definition. Let us take a closer look at the statistics by dictionary editions.

About half of the successful *LDOCE3* consultants depended on the example in deciding on the appropriate sense and they all succeeded in sub-entry search. The example reads *I awoke from my drunken stupor feeling nauseous*. Thanks to its typical contextualisation, their mental association was facilitated, linking *nauseous* there to the one in the original sentence, even though some did not understand the meaning of *stupor*¹¹. In contrast, among the

LDOCE2 users who based their choice on the example, half were misled or trapped. This is because the example under Sense 1 (*nauseous medicine*) provides no particular referential help, and on top of that *nauseous* under Sense 2 translates into the same adjective in Japanese (*kimochi warui*) as the key word in the original sentence¹². The successful *LDOCE2* user who is categorized under “other” mentioned that the *fml/infml* labels led her to the right decision, but in this case they cannot constitute critical factors.

Eight students who used *COBUILD2* did not give any answers. Probably they thought that they were not supposed to do so, finding that the entry exhibited the appropriate sense only. Ten (66.7%) of the *COBUILD1* users relied on the definition for their final decision on the appropriate sense. The percentage is twice as high as that of each version of *LDOCE*, which suggests that the sentential definition is approachable to users. It has to be noted, however, that 40% failed to make proper use of it in determining an appropriate sense. This implies that the defining sentence may not have been so easy to handle (i.e. identify and extract substance and make proper use of it) as it looks. In this particular case success seems to have rested on the understanding of “feel sick” as synonymous to “feel nauseous”, precise and solid enough not to be shaken by the related other sense.

Five subjects across the groups (four were successful at reference) named the original sentence as a reason for their semantic decision. The ability to (re)examine the original context of the look-up item before and during dictionary consultation—in the process of (re)application and (re)integration of dictionary information to the original text—is an important key to success.

Next, the influence of the prior familiarity with the look-up item on reference is investigated. Those students who offered the appropriate or satisfactory translation of the target item in Task 1 (pre-reference translation) are judged to have been familiar with the look-up item; those who gave a wrong or no translation are considered to have lacked such familiarity. Table 3 sets out for each dictionary the relationship between the previous knowledge of the lexical item and sub-entry search.

Table 3 Prior familiarity with the look-up item and sub-entry search

Prior familiarity with look-up item (no. of subjects)	<i>LDOCE3</i> (15 subjects)		<i>LDOCE2</i> (14)		<i>COBUILD2</i> (12)		<i>COBUILD1</i> (15)	
	+	-	+	-	+	-	+	-
	(8)	(7)	(6)	(8)	(7)	(5)	(5)	(10)
Reference success	8 (100%)	5 (71.4%)	4 (66.7%)	3 (37.5%)	N.A.*		4 (80%)	7 (70%)
Reference failure	0 (0%)	2 (28.6%)	2 (33.3%)	5 (62.5%)	N.A.*		1 (20%)	3 (30%)

* Referential success/failure does not apply to the consultation of *COBUILD2*, which exhibits the appropriate sense only.

The “+” and “-” signs indicate possession and lack of pre-consultation familiarity with the target item, respectively. To take the *LDOCE3* group for example by way of explaining what the table represents, eight of them were familiar with (knew or were able to guess) the meaning of *nauseous* before dictionary consultation, and seven lacked such knowledge. All eight students “in the know” (100%) consulted the dictionary successfully for the meaning of the target word by choosing the appropriate sense in Task 2. Out of the seven students who lacked previous knowledge of the word, five (71.4%) were successful at dictionary reference but two (28.6%) were not.

General survey reveals, as expected, that the prior familiarity with the look-up item has a positive effect on dictionary consultation. Although the dictionary is primarily a tool for consulting an unfamiliar meaning of a given lexical item, paradoxically, it is understandable that lack of prior familiarity with the look-up item makes reference more of a burden on the user than when he/she has such knowledge¹³. This manifested itself most clearly in the case of *LDOCE2*, which does not offer any particular referential help or clues, providing only phrase definitions and examples. The dictionary turns out to be hardest on those who lacked previous knowledge of the target word—only 37.5% managed to identify the proper sense, compared with over 70% success rates with the other dictionaries.

5.5 Reference speed

As mentioned at the end of Section 5.3, successful consultation was

performed in a shorter time than unsuccessful consultation with each dictionary. In this test, against expectations, reference of the dictionaries with phrase definitions took longer than those with full-sentence definitions, which should not be taken at face value. This is because, while the test sheet for the *COBUILD* groups provided only the entry on *nauseous*, that for the *LDOCE* groups included the three or four entries between *nausea* and *nauseous*. This was to allow the students to refer to the entry of the noun which is used in the definition of the adjective (look-up word) and they seem to have done it. Repeated below are the mean reference times of successful and unsuccessful consultations with the number of subjects for each dictionary from Table 1.

Table 4 Mean reference times (seconds)

Reference	<i>LDOCE3</i> (15 subjects)	<i>LDOCE2</i> (14)	<i>COBUILD2</i> (12)	<i>COBUILD1</i> (15)
Suc. (no. of subjects)	63.7 (12*)	41.2 (6*)	N.A.	27.9 (10*)
Fail. (/)	113.3 (2)	90.0 (7)	N.A.	36.3 (4)

* One subject from each category is ignored; she changed her mind after entering an appropriate sense number in Task 2.

Focusing on successful references above, I probe the relationship between the reference times and post-reference translations of the target word (success or failure thereof) to see whether there are any correlations. The results are presented in Table 5.

Table 5 Post-reference translation of the look-up item and the times of successful reference (seconds)

Post-ref. translation	<i>LDOCE3</i> (12 subjects)	<i>LDOCE2</i> (6)	<i>COBUILD2</i> (12)	<i>COBUILD1</i> (10)
Suc. (no. of subjects)	70.0 (11)	53 (4)	N.A.	27.9 (10)
Fail. (/)	33.2 (1)	17.5 (2)	N.A.	- (0)

From this data, unfortunately, reliable generalisations cannot be made with the scant numbers or the absence of the subjects who were able to locate the proper sub-entry (Task 2) but unable to translate the look-up word appropriately, based on the elicited information (Task 4). However, it deserves noting

that in both cases of *LDOCE* those successful translators spent much more time on reference; that is, careful examination of dictionary text led to a successful post-reference translation. This makes one suspect that it is largely because of haste that the few subjects failed to finish off their reference successfully by integrating dictionary information into the original text even though they managed to spot the relevant sub-entry.

Next examined is the link between the reference time and the prior familiarity with the look-up item. Table 6 sets out for each dictionary the mean reference times of the subjects with and without the pre-reference familiarity with the look-up word. Once again, only successful references are considered.

Table 6 Prior familiarity with the look-up item and the times of successful reference (seconds)

Prior familiarity with look-up item	<i>LDOCE3</i> (13 subjects)	<i>LDOCE2</i> (7)	<i>COBUILD2</i> (12)	<i>COBUILD1</i> (11)
+ (no. of subjects)	66.9 (7*)	53 (4)	N.A.	26.7 (4)
- ()	60.3 (5)	17.5 (2*)	N.A.	28.7 (6*)

* One subject from each category is ignored; she changed her mind after entering an appropriate sense number in Task 2.

That prior familiarity with the look-up word will have a positive effect on dictionary consultation and also be reflected in the reference time is a plausible hypothesis. However, this is rejected by the statistics from *LDOCE2* and *LDOCE3*, especially the former. The *COBUILD1* case supports the hypothesis but not by a significant margin. Prior familiarity with the look-up item may not exert as much influence on reference speed as on reference success (identification of appropriate sub-entry); there may be other contributing factors which determine reference time.

5.6 Subjects' understanding of definitions

As instructed in the second half of Task 5, students underlined the items in the definitions they had difficulty understanding. The results are tabulated in Table 7. Items are arranged, dictionary by dictionary, in descending order

of the number of the students who marked them. Some mentioned more than one item.

Table 7 Difficult items in definitions

LDOCE3 (15 subjects)		LDOCE2 (14)		COBUILD2 (12)		COBUILD1 (15)	
<i>AmE</i>	4	nausea	4	vomit	10	vomit	10
nausea	1	<i>fml</i>	2			PRED	3
		causing	1			queasy	3
						ADJ QUALIT	1
						QUALIT	1
Total	5	Total	7	Total	10	Total	18

This gives us a general idea of which items caused difficulty¹⁴. However, it should be noted that it may not exhaust all the troublesome items for students. There might have been some who did not indicate any difficult items because of time limitation or out of carelessness. Several students unexpectedly underlined difficult items outside the relevant definition (in the other definition and examples) which are not considered here. Some grammatical abbreviations and labels are included since no prior instruction was given on them. In considering the results on *nausea*, it needs to be taken into account that there might have been some students who did not underline the word even without prior knowledge of it—looking at the entry listed on the same sheet, they may have decided that they now knew the meaning of the noun.

Since the definitions concerned are simple, the students' translation does not deserve much attention. However, a noteworthy pattern was observed among those who translated the *LDOCE* definitions. Out of the 20 subjects who consulted the entry correctly and translated the relevant definition, five (25%) interpreted the verb ending *-ing* of *feeling* and *causing* as nominal (*-koto*), presumably to make it sound like a dictionary definition. Although this may be of no great significance, they need to be taught the general rule of dictionary definitions that the phrase given in such forms is interchangeable with the headword. In this case, the *-ing* ending does not signify a nominal

status; it was employed simply because the headword was an adjective. The most common errors found in the translation of the *COBUILD* definitions resulted from the inability to translate “vomit”.

Table 8 re-presents the averages of the students’ understanding of the definition in question from Table 1 with the averages of the total and the numbers of subjects for each category added.

Table 8 Students’ understanding of the definition in question

Reference	<i>LDOCE3</i>	<i>LDOCE2</i>	<i>COBUILD2</i>	<i>COBUILD1</i>
Suc. (no. of subjects)	73.8% (13)	56.4% (7)	N.A.	84.1% (11)
Failure (/)	12.5% (2)	54.2% (6*)	N.A.	20.0% (4)
Total (/)	65.7% (15)	55.4% (14)	53.3% (12)	67.0% (15)

* One subject did not indicate the rate of her definition understanding.

Here I compare the results of the two versions of each dictionary. The statistics show that those who consulted *LDOCE3* understood the definition better than those who referred to *LDOCE2*, although both definitions are in two-word phrase—“causing nausea” (2nd ed.) and “feeling nausea” (3rd ed.). There will be no problem in establishing that the key factor lies not in the definitions *per se* but in the accompanying examples—the one in *LDOCE3* offers a helpful clue to the users in making sense of the definition while that in *LDOCE2* does not (see 5.4).

On the other hand, it is not so simple to ascertain why the relevant definition in *COBUILD1* was better understood than that in *COBUILD2*. Here are both definitions:

If you feel **nauseous**, you feel sick and as if you are likely to vomit.
(*COBUILD1*)

If you feel **nauseous**, you feel as if you want to vomit. (*COBUILD2*)

The second appears shorter and easier in less complex construction. There seems to be no significant difference in the examples offered as a help to under-

standing the definition, apart from the number—the second edition provides two while the first edition one. Ten subjects from each group mentioned “vomit” as a difficult vocabulary item in the definitions in Task 5. It is plausible that the part “feel sick”, which is unique to *COBUILD1*, offered a clue to understanding, or reduced or compensated for the minus effect of the unfamiliarity with “vomit”. A lesson from this is that verbosity in definition should not always be blamed.

Unlike with *LDOCE3* and *COBUILD1*, the average of *LDOCE2* definition understanding by those successful at sub-entry search is almost the same as that by those unsuccessful. It follows that, while in *LDOCE3* and *COBUILD1* those students who were able to understand the definition well tended to locate the appropriate sense, this does not apply to the consultation of *LDOCE2*. This point will be further discussed at the end of this section in connection with previous knowledge of the look-up item.

The relationship between the students’ post-reference translation of the look-up item and their understanding of the definition is as Table 9 exhibits. Only those students who successfully identified the appropriate sub-entry in Task 2 were considered. Understandably, there is a strong correlation—those successful at the post-reference translation of the look-up word understood the definition much better than those unsuccessful, though the results of *COBUILD1* did not allow comparison.

Table 9 Post-reference translations of the look-up item and understanding of the definition

Post-ref. trans. (no. of subjects)	<i>LDOCE3</i> (13)	<i>LDOCE2</i> (7)	<i>COBUILD2</i> (12)	<i>COBUILD1</i> (11)
Reference success	80% (11)	64% (5)	59% (10)	84.1% (11)
Reference failure	40% (2)	37.5% (2)	25% (2)	– (0)

Lastly, the students’ understanding of the definition is investigated in relation to their prior familiarity with the definiendum. The results set out in Table 10 are quite reasonable—the definition was better understood by those students with previous knowledge of the headword item.

Table 10 Prior familiarity with the look-up item and understanding of the definition

Prior familiarity with look-up item	<i>LDOCE3</i>	<i>LDOCE2</i>	<i>COBUILD2</i>	<i>COBUILD1</i>
+ (no. of subjects)	82.5% (8)	59% (6)	70% (7)	70% (5)
- (/)	46.4% (7)	53.1% (8)	30% (5)	65.5% (10)
Total (/)	65.7% (15)	55.4% (14)	53.3% (12)	67.0% (15)

On closer examination, among the users of *LDOCE2* and *COBUILD1*, there is only a slight discrepancy between the understanding rates of those with and without prior familiarity with the target word—previous knowledge of *nauseous* had much less influence on the comprehension of the definition than in the other dictionaries¹⁵.

Considering the students' understanding of the definition in connection with referential success and prior familiarity with *nauseous*, consulting the adjective in *LDOCE2* seems to have been left largely to chance. When there are no particular contextual clues, grammatical or lexical, in the original context (as in this case), the student has to consult the look-up word in the dictionary by meaning, with or without previous knowledge of the word. However, providing only abrupt phrase definitions and examples, *LDOCE2* does not offer any particular help to the user in referring to the two-sense entry on *nauseous*. This is supported by the evidence that the rate of definition understanding stands around 55%, whether or not they were successful at locating the appropriate sense (Table 8). The information provided is so simple and lacking in clues that even the students with prior familiarity with the look-up word could not take advantage of it in interpreting the definition (Table 10). Consequently, half of the students ended in failure in the effort to identify the relevant sub-entry.

5.7 The impact of dictionary consultation

In order to assess the impact of dictionary use on the subjects' understanding of the target item, a comparison is made between the target word parts of their translations before and after consulting the entry (their trans-

lations in Tasks 1 and 4). The results are set out in Table 11.

Table 11 Students' pre- and post-reference translations

Reference	LDOCE3		LDOCE2		COBUILD2	COBUILD1	
	Success	Failure	Success	Failure	N.A.	Success	Failure
(1st → 2nd trans.) ¹ Right → Right	7 (Refined ² 1)		4	2	7 (Refined ² 1)	4	
↯ → OK							
↯ → Wrong							1
↯ → None	Rev ³ 1						
OK → Right							
↯ → OK							
↯ → Wrong							
↯ → None							
Wrong → Right	2					1, Rev ³ 1	
↯ → OK							
↯ → Wrong	1	1	1	2	2		1
↯ → None		1	1				
None → Right	2		Rev ³ 1	1	3	5	
↯ → OK							
↯ → Wrong				1			2
↯ → None				1			
Subtotal	13	2	7	7	—	11	4
Total	15		14		12	15	

1 The furthest left-hand column indicates the students' performance in translation: their pre-reference translation (Task 1) precedes the arrow and their post-reference translation (Task 4) follows it. The words represent the following:

- Right: appropriate translation
- OK: just satisfactory translation
- Wrong: wrong translation
- None: no translation

2 "Refined" means the post-reference translation was even better than the pre-reference one that was already good enough.

3 "Rev" refers to those who "revised" their decision on the choice of the appropriate sense.

I will illustrate what the table indicates, picking out the *LDOCE3* group's performance. As is shown in the total, 15 subjects used the entry from the dictionary; 12 of them succeeded in their reference by stating the correct

sense number in Task 2, one revised (“Rev”) and succeeded, and two failed. From the top of the list, seven students translated *nauseous* appropriately into Japanese before and after looking up the entry (in Tasks 1 and 4 both, “Right → Right”); one of them gave an even better translation after reference (“Refined”). All of them consulted the entry properly. One gave the appropriate translation before look-up but did not give any after (“Right → None”), although she was successful at sub-entry search. Two gave wrong translations before reference but appropriate translations after (“Wrong → Right”); both succeeded in their look-up. Two translated wrongly before and after look-up (“Wrong → Wrong”); one consulted the entry correctly but the other wrongly. One gave a wrong translation before consultation and no translation after without referring to the entry properly (“Wrong → None”). Two who failed to give any translation before look-up gave the appropriate equivalent after (“None → Right”); each reference was successful.

In an effort to assess the impact of dictionary use on the subjects’ understanding of the target word, I will examine the relationship between their reference (Task 2) and their post-reference translation (Task 4). Analysis is made in terms of successful versus failed sub-entry search and appropriate translation versus wrong or no translation. Although both represent the inability to give a proper translation, wrong translation and no translation should be separately treated. It is important to note that, while the failure to give any translation before look-up can be ascribed to ignorance, the inability to do so after look-up may as well result from a shortage of time or simply negligence on the part of subjects. The results are as shown in Table 12.

Table 12 Students’ post-reference translation

Reference	LDOCE3		LDOCE2		COBUILD2	COBUILD1	
	Success	Failure	Success	Failure	N.A.	Success	Failure
Right. 2nd trans.	11 (73.3%)	0 (0%)	5 (35.7%)	3 (21.4%)	10 (83.3%)	11 (73.3%)	0 (0%)
Wrong/No 2nd trans.	1/1 (13.3%)	1/1 (13.3%)	1/1 (14.3%)	3/1 (28.6%)	2/0 (16.7%)	0/0 (0%)	4/0 (26.7%)
Subtotal	13 (86.7%)	2 (13.3%)	7 (50.0%)	7 (50.0%)	—	11 (73.3%)	4 (26.7%)
Total	15		14		12	15	

Here is how to interpret the table; let us look at the results of the *LDOCE3* group. Among those who used the entry from the dictionary 13 succeeded in their reference and two failed (in Task 2). Eleven of the successful users translated *nauseous* appropriately after consulting the entry (in Task 4). Even though their consultation was successful, two could not make proper use of it in translating the target word: one gave the wrong translation and the other no translation. The two who failed in sub-entry search also failed in translation: one gave the wrong translation and the other no translation. The success/failure distinction does not apply in the consultation of *COBUILD2*, which offers only the proper sense.

I will next take a closer look at the “mismatch” boxes—those who succeeded in their consultation of the entry but failed in translation, and vice versa. The first category covers those who looked up the target word correctly in the attached dictionary entry by choosing the right meaning but could not apply the elicited information properly in their translation effort. There were two such students in the *LDOCE3* group. The one who gave the wrong translation repeated the mistake in both translations—she seems to have made a wrong association between *nauseous* and *noisy* because of the phonetic similarity involving common sounds (her understanding of the definition was 10-1%). The one who gave no translation for the second time translated the target word appropriately for the first time—she did not have time or did not bother to do so for the second time (her understanding of the definition was 89-64%). In the *LDOCE2* group, one repeated the wrong translation both times (her understanding of the definition was 63-37%) and the other who translated wrongly before looking up the entry failed to give any after (her understanding of the definition was 36-11%). There were two cases in the *COBUILD2* group (though they did not have to choose an appropriate entry). Both could not give a proper translation each time (their understanding of the definition was 36-11% each). All these subjects’ failure to give the appropriate post-reference translation of the target word seems to have resulted largely from their inadequate understanding of the relevant definition, apart from the

LDOCE2 user who neglected to give any translation after consulting at the entry.

The *LDOCE2* group contains odd cases where, even though they failed in the look-up, three students succeeded in post-reference translation. Two of them translated *nauseous* properly before looking at the dictionary entry and the other left the target word part blank at the first trial. There are two plausible reasons for this. More probable is the explanation that since both senses of *nauseous* can translate into an identical Japanese adjective (see 5.4) the students translated it on the basis of their understanding of the inappropriate second sense. The other interpretation is that they simply tried a good guess from the information the original sentence offered, like *feeling* and *on the bus*, with success.

In order to gauge the effect of each dictionary on students' understanding of the target word, the look-up item parts of their pre- and post-reference translations are compared from a different perspective. The results are analysed into "Help", "No help" (with or without the right post-consultation translation), and "Adverse effect" and are presented in Table 13.

Table 13 The effect of the dictionaries on students' understanding of the target item

Reference	<i>LDOCE3</i>		<i>LDOCE2</i>		<i>COBUILD2</i>	<i>COBUILD1</i>	
	Success	Failure	Success	Failure	N.A.	Success	Failure
Help	5 (33.3%)		1 (7.1%)	1 (7.1%)	4 (33.3%)	7 (46.7%)	
No help: right 2nd trans.	6 (40.0%)		4 (28.6%)	2 (14.3%)	6 (50.0%)	4 (26.7%)	
No help: wrong/ no 2nd trans.	1/0 (6.7%)	1/1 (13.3%)	1/1 (14.3%)	3/1 (28.6%)	2/0 (16.7%)		3/0 (20.0%)
Adverse effect	1 (6.7%)						1 (6.7%)
Subtotal	13 (86.7%)	2 (13.3%)	7 (50.0%)	7 (50.0%)	—	11 (73.3%)	4 (26.7%)
Total	15		14		12	15	

If the post-reference translation is any tangible improvement on the pre-reference translation, the dictionary can be said to have helped the student in understanding the look-up word (This is what is meant by "Help"). If there is no

such improvement, the dictionary is judged to have offered no help. However, the cases with the right post-consultation translation (“No help: right 2nd trans.”) should be considered separately from those with the wrong or no translation (“No help: wrong/no 2nd trans.”). This is because in the former case the dictionary may have assisted the student modestly in confirming that her pre-consultation understanding of the target word was right. If it did so, the dictionary usefully fulfilled one of its functions (confirming), which may not be so dramatic as imparting knowledge that the user lacked¹⁶. If the post-consultation translation should be worse than the pre-consultation translation in any way, the dictionary is considered to have exerted a deleterious effect (“Adverse effect”)¹⁷.

This analysis may not have done justice to the dictionaries, for example, in conveniently treating *LDOCE3* as exerting a negative effect on the user, who translated the target word appropriately before look-up, succeeded in sub-entry search, but failed to give any translation after (Right → None), where a possibility cannot be ruled out that she did not have enough time or just neglected to do so¹⁸. If this is the case, this is not a problem stemming from the inferiority of the dictionary. Possible events like this should be taken into account in considering cases of “No help: no 2nd trans.” and “Adverse effect” with no post-consultation translation.

Cases like those represented in the “Failure-Help” and “Failure-No help: right 2nd trans.” grids under *LDOCE2* in Table 13 can be dismissed from consideration in estimating the effect of a dictionary because they possibly occurred by pure chance—the dictionary cannot be said to have helped the user who had failed in his/her reference (identification of an appropriate sense), or any help such a user may have obtained is highly incidental.

Considering the nature of this study, which required the students to look up only the item designated by the researcher, regardless of their previous familiarity with the vocabulary items involved, the “Success-Help” and the “Success-No help: right 2nd trans.” grids represent the instances of successful dictionary use—the process from identifying an appropriate sub-entry to

integrating the elicited information to the original context. Therefore, it is legitimate to use these figures to determine the effect of each dictionary on the user's understanding of the look-up word. Assessment based solely on the "Success-Help" grid (to the exclusion of the "Success-No help: right 2nd trans." grid) does not reflect the reality of dictionary consultation required by this study. It does not take into account those successful dictionary consultations where students need the dictionary only for confirmation and will put at a disadvantage the dictionary many users of which just happen to know the designated look-up word, needing only a just-checking type of dictionary consultation (if they ever need any).

The value of a dictionary is hard to determine. Successful dictionary use is the result of a dynamic interplay between the user and the dictionary (the product and indirectly the maker). It is hard to tell which party should be (more) responsible—praised or blamed—for un/successful dictionary uses. In an attempt to estimate the utility of each dictionary in each look-up situation, this study will use as a yardstick the numbers of those subjects who successfully cleared the sequence of these consultational steps—identifying a proper sub-entry, understanding the information provided, extracting the relevant part of the information, and translating the look-up item appropriately by integrating the elicited information into the original context (i.e. those represented in the "Success-Help" and the "Success-No help: right 2nd trans." grids in Table 13). Those dictionaries which guided students all through this process are judged to be good dictionaries. Table 14 presents the rates of successful dictionary use with each dictionary in the look-up situation required by this study. The rates are calculated against the total numbers of subjects and the numbers of successful consultants (those who identified

Table 14 The utility of the dictionaries—the rates of successful dictionary use

	<i>LDOCE3</i>	<i>LDOCE2</i>	<i>COBUILD2</i>	<i>COBUILD1</i>
Out of total	73.3% (11/15)	35.7% (5/14)	83.3% (10/12)	73.3% (11/15)
Out of successful refs.	84.6% (11/13)	71.4% (5/7)	—	100% (11/11)

the appropriate sub-entry in Task 2).

COBUILD2 is rated—the greatest number of students who referred to it successfully completed the consultational process from sub-entry search to information integration—as the most helpful of the four dictionaries, with *LDOCE2*, which lacked any particular user-friendly considerations, dragging far behind. However, the lead of *COBUILD2* should be taken with a grain of salt. Because of its monosemous entry on *nauseous* the dictionary can be considered to have been less demanding on its users—unlike the other dictionaries, it did not confront the subjects with a choice of an appropriate sub-entry. Therefore, it is not unthinkable that this gave the *COBUILD2* users a headstart and enabled them to concentrate on the other consultational requirements, saving up the energy to be directed to sub-entry location, so that the group as a whole fared best in this evaluation.

Over 70% of the subjects who were able to identify the appropriate sub-entry in all dictionaries (except for the monosemous *COBUILD2*) went on to finish off their consultation successfully by translating the look-up word properly. It is noteworthy that all 11 successful *COBUILD1* consultants managed to give the appropriate translation—or the dictionary made it possible.

As discussed above, dictionary evaluation based solely on the “Success-Help” grid in Table 13 does not faithfully represent successful dictionary uses, considering the consultational setting laid by the test. This being so, it has to be remembered that the dictionary is basically a tool for looking up an unknown meaning of a vocabulary item. This test, which checks the subjects’ prior familiarity with the look-up item in Task 1, makes it possible to single out those who lacked such knowledge and observe how well they consulted a dictionary entry. This way, a nearer-genuine dictionary use can be explored and the results will serve as a valid criterion to assess the utility of each dictionary. The same analysis as above is conducted on the subjects without previous knowledge of the look-up item and the results are set out in Table 15 in the same fashion as Table 14.

Table 15 The utility of the dictionaries—the rates of successful dictionary use among the subjects without prior familiarity with the look-up item

	<i>LDOCE3</i>	<i>LDOCE2</i>	<i>COBUILD2</i>	<i>COBUILD1</i>
Out of total	57.1% (4/7)	12.5% (1/8)	60% (3/5)	70% (7/10)
Out of successful refs.	80% (4/5)	33.3% (1/3)	—	100% (7/7)

It is *COBUILD1* that helped most those without previous knowledge of the look-up word in their dictionary use (from sub-entry search to information integration). The fact that the rate (70%) is a slight drop from that in Table 14 (73.3%) suggests that the dictionary assisted those with and without prior familiarity with the target item almost equally, unlike the other dictionaries¹⁹. Once again, *LDOCE2* proves to have been by far the least helpful, helping only one out of the eight subjects.

COBUILD1's triumph continues when we look at the rates of the overall successful dictionary use by those who successfully identified the appropriate sense—without knowing the meaning of the look-up item, seven subjects located the wanted sense and they all translated the target word correctly by extracting and applying the relevant information. This might be ascribed to the defining sentence, one of the dictionary's unique features. In *LDOCE2*, on the other hand, even though three students managed to find the appropriate sub-entry, the dictionary helped only one of them to translate the look-up word. These statistics demonstrates that *LDOCE2* is the least useful to those users who lacked previous knowledge of the look-up item.

5.8 Feedback from subjects

Questions were asked after the test in order to find students' reactions to particular lexicographically significant points and to account for their preferences of the dictionary entries in comparison to the one they did not use in performing the tasks. Two questions were directed to each group. One of them, for the *LDOCE* group, was:

- (1) Both editions define *nauseous* with reference to its noun form *nausea*.

If you do not know the meaning of *nausea*, you have to go to the entry. What do you think of this treatment?

Seventeen (58.6%) reacted negatively and only three (10.3%) positively. Another three said it was neither good nor bad, and six (20.7%) gave no answers. Negative ideas were voiced by using such adjectives as “inconvenient”, “troublesome”, “unkind”, “annoying”. Behind these critical remarks seem to be the following assumptions they harbour rightly: “One look-up at a time should do” and “Those who consult *nauseous* can never be expected to know *nausea*”. In connection with the principle of self-sufficiency of a definition, one suggested as a remedy the inclusion of a bracketed gloss led by an equal sign, like the one put against “VOMIT” in the entry on *nausea* in each edition. All these reflect the user’s point of view. However, it is worth mentioning that there were two who showed understanding for this “inevitable” treatment, hinting at redundancy to be avoided. While finding it inconvenient that one look-up leads them to another, two looked at the positive side, saying that it at least gave them a chance to check the meaning of a word they did not know.

The other question, for the same group, was:

(2) Which entry (definition and example) do you prefer, *LDOCE2*’s or *LDOCE3*’s?

Their reaction was mixed. Only seven made their stance clear while the others were possibly busy answering the first question. The sample is too small to grasp the whole picture but four voted for the third edition and three for the second. Two said *LDOCE3* is easier to understand whereas one said the same thing in favour of *LDOCE2*. One liked the definition of the new edition better, and another pointed out that it gives fuller information, criticising the second edition for not offering full-sentence examples. On the other hand, two praised *LDOCE2* for its conciseness, one of them dismissing *LDOCE3*

as lengthy. One's personal taste seems to manifest itself distinctively in this matter.

Table 16 Students' preference: LDOCE3 vs. LDOCE2

	LDOCE3	LDOCE2
	57.1% (4/7)	42.9% (3/7)
Merits	Easier to understand (2)* Better definition Fuller explanation	Concise (2) Easier to understand
Limitations	Lengthy	No full-sentence examples

* Bracketed is the number of the students who mentioned it when it is more than one.

The *COBUILD* group was first asked:

- (1) The entry on *nauseous* in *COBUILD1* has two senses, while only the first sense is entered in that of *COBUILD2*. Which do you prefer?

The-more-the-better sentiment seems to be basic human nature. Understandably, most (twelve, 44.4%) were in favour of *COBUILD1* and eight (29.6%) of *COBUILD2*. One (3.7%) found advantages in both, and six (22.2%) gave no answers. It is noteworthy that the students supported *COBUILD2* because of its one-sense-only entry. Half found it unnecessary to list two senses as *COBUILD1* does, thinking that the "queasy" and "revolting" senses are so closely related to be discriminated that one is inclusive of the other²⁰. The fact that both can translate into the Japanese adjective *kimotchi warui* may have led them to think this way. The other half preferred *COBUILD2* because of ease of reference—it gave only the context-fit information without bothering them with choice. I wonder what their reaction would have been if the other sense had been the answer. Confronted with an all-L2 dictionary text, novices may have been led to conclude the less look-up load, the better²¹.

As the second question, I repeated the one put to the *LDOCE* group:

- (2) Which entry do you prefer, *COBUILD1*'s or *COBUILD2*'s?

Out of the 12 students who responded, eight (66.7%) preferred *COBUILD1*, three (25%) *COBUILD2*, and one (8.3%) said both. Four liked the first edition because it is easier to understand and two said the same thing with particular reference to its examples. There were opinions that *COBUILD1* is easier to look at and one pointed specifically to the example-demarkation sign “EG”. As reasons for preferring *COBUILD2*, one mentioned its easier definition²² and two its examples (one for their quality and one for their quantity)²³. The one who voted for both liked the new edition for its sentence examples and the old edition for its better definition²⁴. Although *COBUILD1* was much more popular in this case, the subjects’ reaction was mixed and here again the “there is no accounting for tastes” syndrome was in place.

Table 17 Students’ preference: *COBUILD2* vs. *COBUILD1*

	<i>COBUILD2</i>	<i>COBUILD1</i>	Both
	25% (3/12)	66.7% (8/12)	8.3% (1/12)
Merits	Easier definition Better examples More examples Sentential examples	Easier to understand (4)* Better definition (2) Easier examples (2) Easier to look at "EG" sign is useful	
Limitations	Difficult definition Too many examples		

* Bracketed is the number of the students who mentioned it when it is more than one. Some raised more than one point.

5.9 Summary

The findings can be summarised as follows:

- Those subjects who identified the appropriate sub-entry were able to do so quickly, understand the definition well, and make proper use of the elicited information in translating the target word.
- In finalising their decision on the appropriate sub-entry, the *LDOCE* users tended to depend on the example while those of *COBUILD1* depended heavily

on the definition—when there was a choice of the phrase definition and the example, the students preferred the latter; between the sentential definition and the example they went for the former. Overall, the example-oriented sub-entry search was surer than the definition-oriented. Basing the decision on the sentential example resulted in more success than on the phrase example. Examples (especially sentential) which provide sufficient and typical contextualisation help reference and the understanding of the accompanying definition and the headword item.

- The subjects who spent much time on reference succeeded in translating the look-up word, while those who rushed in consultation seem to have failed in the translation. The successful translators understood the definition of the look-up item much better than the unsuccessful ones.
- The ability to (re)evaluate the original context of a look-up item before and during dictionary consultation is an important key to success.
- Prior familiarity with the look-up item has a positive effect on dictionary consultation—sub-entry search, definition understanding, post-reference translation, but not on reference time. Lack of previous knowledge exerts a negative influence particularly on consulting a dictionary which does not offer any special referential help or clues.
- Where the EFL dictionary offers all the riches of information, inexperienced users tend to restrict the range of reference to operable parts. Their referential priority is on the accessibility of information rather than on the quality or relevance thereof. Lamentably, their look-up is predetermined by luck—if a significant referential clue should fall within their reach, their reference is more likely to succeed.
- While the sentential definition is approachable to users, it may not be so easy as it looks to use for sub-entry search. However, once successfully located, the defining sentence securely helps him/her to translate the look-up item.
- One look-up at a time should suffice: a majority of the students do not like the definition which entails an additional look-up—defining a word

with reference to a word of the same derivation (e.g. *nauseous* by means of *nausea*).

- Verbosity in definition is not necessarily detrimental, as long as it leaves room for including footholds to understanding.
- Ignorant that a dictionary definition is often given in such a form that is interchangeable with the headword item, several students wrongly interpreted the *-ing* form in the definition of as nominal.
- There is no accounting for tastes for dictionaries: students' reactions are variegated. Priority seems to be on the quality of being easy to understand and (moderate) quantity of information. Confronted with an all-L2 dictionary text, novices tend to conclude the less look-up load, the better.

6 Conclusion

Despite its limitations, this study based on the unique research and analysis methodology has produced interesting findings and insights into dictionary use (and making). It will be interesting to replicate the study to observe subjects' look-up behaviour longitudinally and to look into various dictionary features with various kinds of lexical items. An interview could be incorporated to detect errors—when they occur and what/who are responsible for them. Even small studies and case studies are revealing and meaningful. They should be gathered, shared, and used to improve lexicography, as was mentioned in the discussion following Professor Hartmann's Asialex lecture (Hartmann 2003).

It was back in 1981 that Béjoint and Cowie pointed out that there was a gap between too sophisticated dictionary design features and students' rudimentary reference skills. More than 20 years later, in spite of dictionary makers' endeavours to make their products better and more accessible, the situation seems to remain more or less unchanged—dictionaries are not fully utilised. Education on dictionaries and dictionary use should figure in a major way in an effort to narrow the gap.

Notes

- * I would like to express my special thanks to Professor Reinhard Hartmann and Professor Richard Murto for their valuable comments and help in preparing this paper.
- 1 Attached are the original test sheets, whose directions and questions are given in Japanese.
 - 2 These dictionaries were chosen for the following reasons and considerations. For the purpose of this study, just a few representative and well contrasted EFL dictionaries were necessary. To give it a historical perspective, two sets of two consecutive editions were appropriate. The third and fourth generation dictionaries (cf. Minamide 1998, Cowie 1999) were considered for use because the latter were the latest when the study was conducted. Since my research interest is in the current dictionaries and the study is partially educationally oriented, it was decided to be sensible that students should be exposed to dictionaries which describe up-to-date English. A final choice was thus made of the *LDOCEs* and the *COBUILDs*, which include some of the remarkable features representing lexicographical traditions and innovations.
 - 3 According to the questionnaire given in the first meeting, one subject studied in Oregon for a year during her high school years. Another was born and raised in New York for four years. The other spent one and a half years of her junior high school years in New York.
 - 4 In the same questionnaire, two subjects of Class A named "Oxford" English-English dictionary with no further specifications. One Class B student mentioned "Dai ei-ei jiten" (literally "large English-English dictionary"). Although *Idiomatic and Syntactic English Dictionary* has been on sale under the name of *Shin Ei-ei Dai Jiten* in Japan, it was not clear which dictionary she referred to.
 - 5 103 on the counter corresponded to 60 seconds.
 - 6 This was also, from an educational consideration, designed to keep students at the task until time is up.
 - 7 As a very minor alteration, "(1st time)" and "(2nd time)" were inserted into the instructions for Tasks 1 and 5, respectively.
 - 8 The thirteen sentences I used for research and education are as follows. The look-up words are underlined:

Test 1: I often feel nauseous on the bus.

Test 2: She stifled a yawn.

Test 3: She squeezed another book into the box.

Test 4: Traffic moves very slowly at peak hours.

Test 5: They submitted their report to us.

Test 6: When I phoned to confirm my flight, I was told there was a strike at the

airport.

Test 7: I'll miss you terribly when you go away.

Test 8: It can't be helped.

Test 9: These apples have tough skin.

Test 10: The minister tendered his resignation.

Test 11: He has no idea how to work that machine.

Test 12: I always feel at home in France when I go there on holiday.

Test 13: A 1970s fashion revival is sweeping Europe.

- 9 It would have been possible to time how long it took the subjects to judge whether or not the only sense entered corresponded to that of the target word.
- 10 Here again successful reference is considered separately from failed. Hereafter those who made the right choice of meaning in the end are counted as successful users, even if they changed their mind on the way, unless otherwise indicated.
- 11 In response to Task 5, three of them underlined the noun, while they were supposed to underline only difficult items within the definition in question.
- 12 Compared to the phrase example under Sense 1, the sentential example under Sense 2 is easier to understand which allows the user to identify himself/herself with "me" therein: *Violence in films makes me nauseous*. It also appears more approachable than the phrase definitions provided under both senses. Especially under the circumstances where both senses translate into an identical L1 adjective, it is probable that the beginning students were attracted to and utilised what they found readily accessible and usable in the information offered in trying to fill the lexical gap in the original sentence.
- 13 To put it differently, prior familiarity with a look-up item has a positive effect on dictionary consultation, or alleviates look-up loads. This provides an insight into considering and designing dictionary use education, especially of MLDs (monolingual learners' dictionaries).
- 14 This will provide lexicographers with a hint about defining vocabulary compilation for an EFL dictionary.
- 15 As discussed above, the part "feel sick" in the *COBUILD1* definition might have helped to offset the disadvantage of unfamiliarity with *nauseous*.
- 16 How often and how many users turn to a dictionary for confirmation in real-life? is a question that is outside the scope of this study.
- 17 The illustration of each category by means of the descriptions used in Table 11 is this:

Help: Right → Right (Refined), OK → Right, OK → OK (Refined), Wrong → Right, Wrong → OK, None → Right, None → OK

No help: right 2nd trans.: Right → Right, OK → OK

No help: wrong/no 2nd trans.: Wrong → Wrong, Wrong → None, None → Wrong, None → None

Adverse effect: Right → OK, Right → Wrong, Right → None, OK → Wrong, OK → None

18 Another possibility that cannot be denied is that dictionary consultation served to shake her confidence in her pre-consultation understanding of the target word.

19 Here are the statistics on successful dictionary use of those who knew the look-up item before consulting the entry. It is evident that such knowledge greatly facilitated dictionary use.

Table 18 The utility of the dictionaries—the rates of successful dictionary use among the subjects with prior familiarity with the look-up item

	<i>LDOCE3</i>	<i>LDOCE2</i>	<i>COBUILD2</i>	<i>COBUILD1</i>
Out of total	87.5% (7/8)	66.7% (4/6)	100% (7/7)	80% (4/5)
Out of successful refs.	87.5% (7/8)	100% (4/4)	—	100% (4/4)

20 Although taking a the-more-information-the-better stance, one *COBUILD1* voter pointed out that sense discrimination should not be too fine.

21 A similar inclination was observed among senior high school students. When I provided them with phrase definitions and sentence definitions from EFL dictionaries (*CULD*, *COBUILD1*, *LDOCE2*, and *OALD4*), some students preferred the former because they are shorter, requiring less effort. Too much should not be given—this provides an important insight for designing MLDs and training on dictionary use for beginning students.

22 One who liked *COBUILD1* better found the second edition's definition difficult.

23 On the contrary, one in favour of *COBUILD1* said that *COBUILD2* gives too many examples.

24 One agrees to the reason for her preference of *COBUILD1*.

References

1 Dictionaries

- Chambers Universal Learners' Dictionary*. International Students' ed. (*CULD*). 1980. Edinburgh: Chambers.
- Collins COBUILD English Language Dictionary*. 1st ed. (*COBUILD1*). 1987. London and Glasgow: Collins.
- Collins COBUILD English Dictionary*. 2nd ed. (*COBUILD2*). 1995. London: HarperCollins.
- Idiomatic and Syntactic English Dictionary (ISED)*. 1942. Tokyo: Kaitakusha. (『新英英大辞典』語学教育研究所版。1942。開拓社。)
- Longman Dictionary of Contemporary English*. 2nd ed. (*LDOCE2*). 1987. Harlow: Longman.
- Longman Dictionary of Contemporary English*. 3rd ed. (*LDOCE3*). 1995. Harlow: Longman.
- Oxford Advanced Learner's Dictionary*. 4th ed. (*OALD4*). 1989. Oxford: OUP.

2 Others

- Béjoint, Henri.** 1981. "The Foreign Student's Use of Monolingual English Dictionaries: A Study of Language Needs and Reference Skills", *Applied Linguistics* 2. 207-222.
- Cowie, Anthony, P.** 1981. "Lexicography and its Pedagogic Applications: An Introduction", *Applied Linguistics* 2. 203-206.
- Cowie, Anthony, P.** 1999. *English Dictionaries for Foreign Learners: A History*. Oxford: Clarendon Press.
- Hartmann, R. R. K.** 1989. "Sociology of the Dictionary User: Hypotheses and Empirical Studies", in Eds. Hausmann, F. J., et al. *Wörterbücher/Dictionary/Dictionnaires*. Vol. 1 Berlin and New York: Walter de Gruyter. 102-111.
- Hartmann, R. R. K.** 2001. *Teaching and Researching Lexicography*. Harlow: Pearson Education.
- Hartmann, R. R. K.** 2003. "Why Lexicography Needs a Strong Academic Foundation", in Eds. Murata, Minoru, et al. *Asialex '03 Proceedings. Dictionaries and Language Learning: How can Dictionaries Help Human & Machine Learning?* The Asian Association for Lexicography. 42-48.
- Hartmann, R. R. K., ed.** 1999. *Dictionaries in Language Learning. Recommendations, National Reports and Thematic Reports from the TNP Sub-Project 9: Dictionaries*. URL: www.fu-berlin.de/elc/TNPproducts/SP9dossier.doc
- 小島義郎他. 1989. 『COBUILD の分析』 *LEXICON* 18. 39-154. (Kojima, Yoshiro, et al. 1989. "An Analysis of Collins COBUILD English Language Dictionary", *LEXICON* 18. 39-154.)
- Masuda, Hideo, et al.** 1997. "An Analysis of the Collins COBUILD English Dictionary, New Edition", *LEXICON* 27. 18-73.
- 南出康世. 1998. 『英語の辞書と辞書学』大修館. (Minamide, Kosei. 1998. *English Dictionaries and Lexicography*. Tokyo: Taishukan.)
- Nesi, Hilary.** 1999. "The Specification of Dictionary Reference Skills in Higher Education", in Ed. Hartmann, R. R. K. 53-67.
- Scholfield, Phil.** 1982. "Using the English Dictionary for Comprehension", *TESOL Quarterly* 16. 185-94.
- Scholfield, Phil.** 1999. "Dictionary Use in Reception", *International Journal of Lexicography*. Vol. 12. No. 1. 13-34.
- 清水あつ子他. 1989. 『Longman Dictionary of Contemporary English (New Edition) の分析 (1)』 *LEXICON* 18. 155-239. (Shimizu, Atsuko, et al. 1989. "An Analysis of Longman Dictionary of Contemporary English, New Edition (Part 1)", *LEXICON* 18. 155-239.)
- 清水あつ子他. 1990. 『Longman Dictionary of Contemporary English (New Edition) の分析 (2)』 *LEXICON* 19. 30-171. (Shimizu, Atsuko, et al. 1990. "An Analysis of Longman Dictionary of Contemporary English, New Edition (Part 2)", *LEXICON* 19. 30-171.)
- Urata, Kazuyuki, et al.** 1999. "An Analysis of Longman Dictionary of Contemporary English, Third Edition", *LEXICON* 29. 66-95.

- Yamada, Shigeru.** 1996. "Instruction in the Use of English Monolingual Learners' Dictionaries for Comprehension", *Exeter Working Papers in English Language Studies* 1. 104-116.
- Yamada, Shigeru.** 2001. "The Abilities Needed for EFL Dictionary Use and How to Teach Them", *Studies in Lexicography*. Vol. 11. No. 1 Center for Linguistic Informatics Development, Yonsei University. 113-125.

Appendices

Appendix 1: Test sheets for the pilot study

1.1 LDOCE sheet

as·sign·ment /ə'saɪnmənt/ *n* 1 [C] a piece of work that is given to someone as part of their job, or that a student is asked to do: *a history assignment* | *on an assignment* *Joanna's going to Italy on a special assignment for her newspaper.* 2 [U] the act of giving people particular jobs to do: *the assignment of chores*

as·sign·ment /ə'saɪnmənt/ *n* 1 [C] a duty or piece of work that is given to a particular person: *She's going to India on a special assignment for her newspaper.* | *His assignment was to follow the spy.* 2 [U (of)] the act of assigning: *the assignment of the chores*

組 番号 氏名 _____

1 次の英文を日本語に訳しなさい。

[Student to professor]

I'd like to talk to you about the assignment you gave us today.

(1回目)

(2回目)

2 左の辞書の項目の中で、上の英文の単語の意味に該当する語義番号を書きなさい。その際テープのCOUNTER RESETボタンを押してから、PLAYボタンを押し検索を始め、終わったらSTOPボタンを押し、所要時間をカウンターの番号で書きなさい。

語義番号： _____ [] 所要時間： _____

* 後で語義番号の選択の間違いに気がついた場合は、語義番号、所要時間を斜線で消し、正しい番号を[]に補い、作業を続けなさい。

3 上の語義が該当すると思うに至った過程（自分のこの場合の辞書の引き方）を段階毎に述べなさい。

4 1の英文をもう一度訳しなさい。前回と全く同じ場合は「同上」と記す。

5 2で選んだ語義の定義を日本語に訳しなさい。記号、略語を含め、わからない箇所は左の辞書の項目に直接下線を引き、訳さなくてもよい。

6 担当者の模範訳に照らし、自分の定義の理解度を自己評価しなさい。

- | | | |
|---------|----------|-------------|
| a 完璧な理解 | b 半分位の理解 | c 全くわかっていない |
| d aとbの間 | e bとcの間 | f その他 |

1.2 COBUILD sheet

assignment /ə'saɪnmənt/ **assignments**

◆◆◇◇◇

N-COUNT

1 An **assignment** is a task or piece of work that you are given to do, especially as part of your job or studies. *The assessment for the course involves written assignments and practical tests.*

2 You can refer to someone being given a particular task or job as their **assignment** to the task or job. *An Australian division scheduled for assignment to Greece was ordered to remain in Egypt... I hardly ever take photographs except on assignment.*

N-UNCOUNT:
of N 10 N**assignment** /ə'saɪnmənt/, **assignments**. 1 An

assignment is 1.1 a particular task that you are officially given to do, especially as part of your job.

N COUNT

eg My first major assignment as a reporter was to cover a large-scale riot... This would be a challenging assignment. 1.2 a piece of academic work given to students. *eg The course has heavy reading assignments but a flexible assessment system.*

N COUNT

2 The **assignment** of a person to do something or to go somewhere is the act of assigning them to it. *eg The original uproar about his assignment to the case has died down.*

N UNCOUNT:
WITH POSS, IF +
PREP THEN to
= appointment

組 番号 氏名

1 次の英文を日本語に訳しなさい。

[Student to professor]

I'd like to talk to you about the assignment you gave us today.

(1回目)

(2回目)

2 左の辞書の項目の中で、上の英文の単語の意味に該当する語義番号を書きなさい。その際テープのCOUNTER RESETボタンを押してから、PLAYボタンを押して検索を始め、終わったらSTOPボタンを押し、所要時間をカウンターの番号で書きなさい。

語義番号： _____ [] 所要時間： _____

* 後で語義番号の選択の間違いに気がついた場合は、語義番号、所要時間を斜線で消し、正しい番号を[]に補い、作業を続けなさい。

3 上の語義が該当すると思うに至った過程（自分のこの場合の辞書の引き方）を段階毎に述べなさい。

4 1の英文をもう一度訳しなさい。前回と全く同じ場合は「同上」と記す。

5 2で選んだ語義の定義を日本語に訳しなさい。記号、略語を含め、わからない箇所は左の辞書の項目に直接下線を引き、訳さなくてもよい。

6 担当者の模範訳に照らし、自分の定義の理解度を自己評価しなさい。

- a 完璧な理解 b 半分位の理解 c 全くわかっていない
 d aとbの間 e bとcの間 f その他

Appendix 2: Final test sheets

2.1 LDOCE sheet

nau-se-a /'nɔ:ziə, -siə/ n [U] *formal* the feeling that you have when you think you are going to vomit (=bring food up from your stomach through your mouth): *Early pregnancy is often accompanied by nausea.* —see also AD NAUSEAM

nau-se-ate /'nɔ:ziət, -si-/ v [T] to make someone feel NAUSEA: *Even clear fluids were making him feel nauseated.* | *It nauseates me the way Keith bullies you.*

nau-se-a-ting /'nɔ:ziɪtɪŋ, -si-/ adj 1 making you feel NAUSEA: *In summer the smell of the farmyard was nauseating.* 2 making you feel angry: *It's nauseating how the coach always picks his favorites.* —compare DISGUSTING —nauseatingly adv

nau-se-ous /'nɔ:ziəs, -siəs/ adj 1 especially AmE feeling nausea: *I awoke from my drunken stupor feeling nauseous.* 2 formal making you feel NAUSEA: *the nauseous stench of the durian fruit* —nauseously adv —nauseousness n [U]

nau-se-a /'nɔ:ziə, -siə/ n [U] *fml* a feeling of sickness and desire to vomit (=to throw up the contents of the stomach through the mouth): *Early pregnancy is often accompanied by nausea.* | *Do you experience any nausea?*

nau-se-ate /'nɔ:ziət, -si-, -zi-, -zi-/ v [T] to cause to feel nausea; SICKEN: *a nauseating smell* | (fig.) *The way he shouts at his wife nauseates me.* —atingly adv

nau-se-ous /'nɔ:ziəs, -siəs, -ziəs, -fəs/ adj 1 *fml* causing nausea: *nauseous medicine* 2 *infml, esp. AmE* feeling great distaste; nauseated: *Violence in films makes me nauseous.* —ly adv —ness n [U]

組 _____ 番号 _____ 氏名 _____

1 次の英文を日本語に訳しなさい (1回目)。

I often feel nauseous on the bus.

(1回目)

(2回目)

2 左の辞書の項目の中で、上の英文の単語の意味に該当する語義番号を書きなさい。テープのCOUNTER RESETボタンを押してから、PLAYボタンを押して検索を始め、終わったらSTOPボタンを押し、所要時間をカウンターの番号で書くこと。

語義番号： _____ [_____] 所要時間： _____

* 後で語義番号の選択の誤りに気がついた場合は、語義番号、所要時間を斜線で消し、正しい番号を [] に記入してから、作業を続けなさい。

3.1 なぜ上で選んだ語義が該当すると思うか、判断の根拠、手掛かりを書きなさい。

3.2 上で選んだ語義が該当すると思うに至った過程、手順（自分のこの場合における辞書の引き方）を段階毎に述べなさい。

4 1の英文をもう1度訳しなさい (2回目)。1回目と全く同じ場合は「同上」とのみ記すこと。

5 2で選んだ、左の辞書の語義の定義を日本語に訳しなさい。記号、略語を含め、わからない箇所は左の辞書の項目に直接下線を引くこと。

6 担当者の模範訳に照らし、自分の定義の理解度を自己評価しなさい。

a 100% b 99~90% c 89~64% d 63~37% e 36~11% f 10~1% g 0%

2.2 COBUILD sheet

nauseous /nɔːziəs, AM -ʒəs/. If you feel **nauseous**, you feel as if you want to vomit. *If the patient is poorly nourished, the drugs make them feel nauseous... A nauseous wave of pain broke over her.* ADJ-GRADED
=QUEASY

nauseous /nɔːziəs, -siəs/. 1 If you feel **nauseous**, you feel sick and as if you are likely to vomit. *eg I felt dizzy and nauseous.* ADJ QUALIT:
PRED
= queasy
2 Something that is **nauseous** is unpleasant and causes you to feel strong feelings of disgust or dislike. *eg ...the nauseous ugliness of the nightmare.* ADJ QUALIT
= revolting

組 _____ 番号 _____ 氏名 _____

- 1 次の英文を日本語に訳しなさい (1回目)。

I often feel nauseous on the bus.

(1回目)

(2回目)

- 2 左の辞書の項目の中で、上の英文の単語の意味に該当する語義番号を書きなさい。テープのCOUNTER RESETボタンを押してから、PLAYボタンを押し検索を始め、終わったらSTOPボタンを押し、所要時間をカウンターの番号で書くこと。

語義番号： _____ [_____] 所要時間： _____

* 後で語義番号の選択の誤りに気がついた場合は、語義番号、所要時間を斜線で消し、正しい番号を [] に記入してから、作業を続けなさい。

- 3.1 なぜ上で選んだ語義が該当すると思うか、判断の根拠、手掛かりを書きなさい。

- 3.2 上で選んだ語義が該当すると思うに至った過程、手順（自分のこの場合における辞書の引き方）を段階毎に述べなさい。

- 4 1の英文をもう1度訳しなさい (2回目)。1回目と全く同じ場合は「同上」とのみ記すこと。

- 5 2で選んだ、左の辞書の語義の定義を日本語に訳しなさい。記号、略語を含め、わからない箇所は左の辞書の項目に直接下線を引くこと。

- 6 担当者の模範訳に照らし、自分の定義の理解度を自己評価しなさい。

a 100% b 99~90% c 89~64% d 63~37% e 36~11% f 10~1% g 0%