1. Introduction

It can be said that the *OCDSE* is the first *pedagogically oriented* electronic corpus-based English collocations dictionary available in the form of a paper dictionary. According to Nuccorini (2003: 373), the first English dictionary with the word *collocation* in its title is Douglas-Kozlowska and Dzięzanowska’s *Selected English Collocations* (1982). The collocations in this dictionary were all drawn from a collection of British texts dating from after 1960 (p. 11). The first collocations dictionary based on an electronic corpus is *A Dictionary of English Collocations* (1994) compiled by Göran Kjellmer. It is a three-volume dictionary based on the 1-million-word Brown Corpus, and includes not only what we generally regard collocations, but also lexical bundles. This is a scholarly work, but the very small size of the corpus on which it is based severely limits its pedagogical value. *The Collins Cobuild English Collocations on CD-ROM* (1995) is a CD-ROM for teachers and advanced learners that provides
access to 10,000 words that appear more than 500 times in the Bank of English (of 200 million words) and shows their most frequent twenty collocates. Real examples extracted from the corpus are shown in the KWIC (Key Word In Context) format and can be also shown in full sentence.

There has been no firmly established structure of a collocations dictionary so far, but as Nuccorini (2003: 367) says, there is one common feature:

English monolingual collocational dictionaries agree on one point: they are meant for encoding purposes and are consistently addressed to advanced learners and translators. This means that both the macro- and micro-structure are devised and organised to help the user write in English. Their shared purpose remains a unifying element even though the types of headwords, of collocates and of the information given are quite different in each dictionary.

The back cover of the OCDSE says that “Learning how words can be combined in English will allow you to say what you mean — and say it more naturally”. A collocations dictionary can help its users to produce not only correct, but also natural texts. Also, in the process of selection of a most appropriate word, users may be able to clarify or develop or even reshape the idea that they want to express. The LTP Dictionary of Selected Collocations (1997) edited by Jimmie Hill and Michael Lewis, another EFL collocations dictionary, therefore strongly encourages users to browse through the entry, explaining a beneficial effect of browsing or scanning as follows in the introduction (p. 9):

Browsing

Apart from looking for the collocation which says exactly what you mean, you can use an entry, particularly one of the longer ones, to help you to re-activate lots of half-remembered combinations. Use the entry in this way when you are preparing an essay or a talk.

The present paper aims to assess the OCDSE’s performance as a corpus-based EFL dictionary of collocations from the foreign learner’s point of view. After giving an overview of the OCDSE, the focus will be on its coverage in order to find out what kinds of collocations or word-combinations the OCDSE actually regards as of practical use. I will take a comparative approach in the analysis of coverage, and the following three dictionaries are used for comparison: the BBI Dictionary of English Word Combinations (1997) (hereafter abbreviated as BBI), the Kenkyusha Dictionary of English Collocations (1995) (hereafter abbreviated as KDEC), and the Kenkyusha Luminous Japanese-English Dictionary (2001) (hereafter abbreviated KLJED). (For more information about each dictionary see Section 7.)

2. The definition of ‘collocation’ in the OCDSE

There has been so far no agreement on the definition of ‘collocation’ among linguists, and the compilers of the OCDSE explain collocation in the section called ‘What is collocation?’ in ‘Introduction’ as follows: “Collocation is the way words combine in a language to produce natural-sounding speech and writing” (vii). This linguistically loose definition of collocation reflects clearly the coherent editorial policy of this dictionary. It is not a linguistically precise definition of collocation, which general users would not understand or care about, but users’ needs for production that forms the basis of judgement as to what collocations or word-combinations are to be included. In their paper explaining the principles and practice of the compilation of the OCDSE, Lea and Runcie (2002: 819) argue for “a pragmatic and fairly wide-ranging definition of collocation for the purposes of the dictionary, based firmly on the needs of the user”.

Although there is no theoretical account of collocation provided, it is explained that they see collocation on a cline with ‘free’ word-combination such as see a man|car|book at one end, and with ‘fixed and idiomatic’ word-combination such as see the wood for the trees at another, and that anything that comes between them can be seen as collocation. They divide collocation further into three categories: weak collocation, medium-strength collocation, and strong collocation. See a film is given as an example of a ‘weak’ collocation, while see danger|reason|the point are examples of ‘strong’ collocations (vii). An example of a ‘medium-strength’ collocation is see a
doctor, although it is NOT included at the entry for doctor or dentist. According to Hill (2000: 64), learners are generally unsuccessful in making efficient use of their knowledge of independent lexical items, especially in the area of 'medium-strength' collocations. The OCDSE therefore attaches pedagogical significance to 'medium-strength' collocations and claims to give them a good coverage.

3. The inclusion policies
Under this broad definition of collocation, the OCDSE basically tries to include lexical strings of words they believe to be pedagogically significant, rather than collocations in the linguistically precise sense. In ‘Which collocations are included in this dictionary?’ in ‘Introduction’ three basic questions asked as the criteria are given: “Is this a typical use of language? Might a user of this dictionary want to express this idea? Would they look up this entry to find out how?” (viii). In other words, what is most important as a criterion for inclusion is whether the dictionary provides what intermediate to advanced learners of English would need in order to produce natural or native-like English. More specifically, the dictionary claims to give an extensive coverage of collocations from 'weak' to 'strong' and exclude free word-combinations and idioms (viii). With some 'strong collocations', whose meaning may be (partially) opaque, a gloss may be provided to clarify the meaning.

‘Frequency’ is a major, deciding factor in inclusion, as well as ‘usefulness to users’, and these two factors are considered interdependent. Collocations native speakers produce frequently are considered to be the ones users are also likely to need to express their idea in more natural English, and the most frequent collocations turned out to be ‘medium-strength’ collocations (Lea and Runcie 2002: 822–25). Accordingly, word-combinations which are not strictly collocations may also be included if they have high enough frequency in the corpus data.

The emphasis on the typical led to a fairly inclusive approach in terms of the perceived 'strength' of collocations . . . . The dictionary does include a number of items that are so fixed that they might properly be called compounds: examples are grey area and learning curve; . . . the learner is not to know the most appropriate expression is so idiomatic that it qualifies as a compound, and the distinction may not therefore be very useful.

(Lea and Runcie 2002: 821)

In their view, ‘typicality’ is judged by ‘frequency’, not by ‘lexicographer’s intuition’, and Lea and Runcie (2002: 824) conclude that “What learners of English really need to know are the collocations that native speakers actually use, in the course of writing and speaking on any subject but collocation; not necessarily what ‘springs to mind’ when a native-speaker is asked to name a collocation”.

British English is chosen as the standard, and ‘moderately formal language’, which is explained as “the language of essay and report writing, and formal letters” (viii–ix), is considered the most appropriate style level for the potential users. Collocations from some technical fields and spoken language are also included and labeled as such. Japanese learners of English would very much appreciate information about American English, however.

4. Methodology of selection of material
4.1. The headword list
The OCDSE claims to contain 9,000 headwords, most of which are commonest words and considered to be familiar to upper-intermediate learners (x). The headword list, which consists of nouns, verbs, and adjectives, was drawn up based on the Oxford Wordpower Dictionary 3), another Oxford product for intermediate learners, and the Oxford Advanced Learner's Dictionary 4), one of the best-known learners' dictionaries (Lea and Runcie 2002: 825). How to draw up a headword list that meets users' needs would require a full, in-depth discussion, so that I will not go further than mention some problems in this paper. First, although it is stated that very common words such as make and do ought to be excluded (Lea and Runcie 2002: 826), there seems to be some inconsistency in the inclusion. While make, do, have, and give are not included, another delexical verb take is given an entry, albeit very short. Treatment of verbs that usually deserve a long entry in a general dictionary, and that
generate a number of phrasal verbs does not seem quite systematic either, as come, get and set are not included, but go, bring, keep, and put are. Second, the headword list contains some culture-specific items. For example, are foreign learners of English likely to talk about or write about A level? Even when they need to do so, they can always refer to the entries for exam and test. Among basketball, football, and hockey, it is of course football that boasts the longest entry, and the entry for cricket is longer than that of basketball and that of hockey, even longer than that of tennis. Since some collocates found at these entries are common to quite a few sports and there is a special page for Sport, the relevant collocations could have been presented more efficiently. Third, some headwords such as bailiff, focal point, firing squad and physique appear to be too difficult or technical for the target users. There seems to be still room for improvement in terms of the headword list.

4.2. Collocations

Basically, all the collocations in the OCDSE are derived from corpus data. Lea and Runcie (2002: 820) explain that lexicographers checked the frequent collocates of a given word in the corpus, going through concordance lines and using MI-score and T-score, and that “A few of the collocations included in the Oxford Collocations Dictionary have as few as three citations in the 100 million word BNC” (825). Lexicographers are also supported by a collocation-extract software program developed by the team (820). As for those new words in computing and the like, the Internet was used as a supplementary source (viii).

However, it is not the case that collocations of high frequency were blindly taken from the corpus and presented in the entries. Lexicographers’ intuition plays an important role in compiling lists of typical words to go with the headword, but not those of frequent words. Rare though it was, the corpus data was supplemented by lexicographers’ intuition when the data did not include items that would be expected to be in. Some words were judged to be excluded from the lists if they appeared highly frequently merely because they were semantically related to the popular topics of the time when the data was collected (Lea and Runcie 2002: 824).

5. The entry structure

Now that editorial policies have been explained, the entry structure will be looked at in this section. Information contained in the entry deals almost exclusively with the combinatorial behaviour of the headword. Figure 1 is the entry for damage, which may be considered as a typical noun entry. The headword is presented in bold, together with its part of speech in italics. (There is no information about the pronunciation provided.) Different parts of speech of a word are given separate entries, so that the entry for the verb damage follows immediately after the one for the noun. Within an entry, different word senses are given separate numbered sections, each with a short definition or pointer to the appropriate meaning: “1 harm/injury” and “2 damages money you can claim from sb”. (However, when there is only one word sense, no definition or pointer is provided.) The plural in bold face in the definition for the sense 2 shows that the headword takes the plural form in this sense. Collocates are grouped and presented together according to their parts of speech. At the noun entry, adjective collocates are first given at the section marked with •ADJ. Then, verb collocates that take damage as its object are given at • VERB + DAMAGE, and verb collocates that take damage as its subject at • DAMAGE + VERB. • PREP. represents prepositions and contains ones that are used with damage. Usually at the end of the section are given other miscellaneous phrases (• PHRASES). Within each section, collocates are further, semantically or thematically grouped into subsections, which are separated by a vertical bar and are claimed to be arranged “in an order that tries to be as intuitive as possible” (x). Within a subsection, collocates are arranged in alphabetical order. For example, the first string of the adjective section goes: “considerable, enormous, great, serious, severe, substantial, untold | minor, slight | extensive, widespread | irreparable, irreversible, lasting, long-term, permanent”. Phrase or sentence examples are given to some collocate subsection: “damages incurred by the unfairly sacked workers”, “He decided to sue the company for damages”. However, the criteria for which word senses to give examples are not explained.
damage noun

1 harm/injury
• ADJ. considerable, enormous, great, serious, severe, substantial, untold | minor, slight | extensive, widespread | irreparable, irreversible, lasting, long-term, permanent

The incident did permanent damage to relations between the two countries. | criminal, malicious, wilful

He was prosecuted for criminal damage to a vehicle.
| accidental

The insurance policy covers the building for accidental damage. | emotional, environmental, mechanical, psychological, structural | brain

She suffered serious brain damage at birth.

• VERB + DAMAGE cause, do, inflict

The earthquake caused widespread damage to property. | They inflicted severe psychological damage on their opponents. | suffer | repair | prevent

• PREP. by

The palace suffered extensive damage by fire in 1825. ~ from Crops are sprayed with chemicals to prevent damage from insects. ~ to lasting damage to the environment

• PHRASES the cost/value of the damage

The cost of the damage is estimated at around $2 billion. | the extent of the damage

At the moment it is difficult to assess the extent of the damage.

2 damages money you can claim from sb
• ADJ. heavy, substantial

• VERB + DAMAGES incur, suffer

damages incurred by the unfairly sacked workers | claim, seek, sue (sb) for He decided to sue the company for damages. | assess

The court will assess the damages. | apportion, award (sb) | pay (sb) | obtain, receive, recover, win | be liable for, be liable in (law) If goods are lost in transit, the carrier will be liable for damages.

• DAMAGES + NOUN action

A woman is to bring a civil damages action against the men she alleges murdered her son. | award

• PREP. in

They are claiming £65 million in damages. | for He received damages for personal injury. ~ of She was awarded damages of £90 000.

• PHRASES an action/a claim for damages

The judge upheld her claim for damages against her former employer.

Figure 1: the entry for damage (noun) from the OCDSE

6. The types of collocations covered

6.1. Formal categories

As is mentioned above, collocates are grouped together according to their parts of speech. The types of collocation dealt with in the OCDSE are shown with some examples blow:

1. Under noun headword:

adjective + noun collocations
brief conversation, mutual cooperation
quantifier + noun collocations
an element of compulsion, a cup of coffee
noun + verb collocations
controversy arises, a competition takes place
verb + noun collocations
carry on a conversation, go into competition
noun + noun collocations
credit account, credit limit, crisis management
preposition + noun collocations
on credit, under curfew
noun + preposition collocations
a crime against . . . , competition between . . .

phrases
a letter of credit, the scene of the crime

2. Under verb headword:

adverb + verb collocations
fully comprehend, exactly coincide
verb + verb collocations
fail to comprehend, refuse to compromise
verb + preposition collocations
correspond with . . . , compromise on . . .

phrases
curse the day, pick and choose

3. Under adjective headword:

verb + adjective collocations
go crazy, remain calm
adverb + adjective collocations
totally convincing, absolutely correct

adjective + preposition collocations
be crazy at . . . , be curious about . . .

phrases
fed and clothed, chilled to the bone

Unlike some other collocations dictionaries such as the BBI and the KDEC, the OCDSE includes no grammatical construction patterns. It may be wise to exclude descriptions of grammatical behaviour of a headword, regarding it as being outside the scope of a collocations dictionary,
and focuses on lexical behavior of a headword, especially when the dictionary is subject to severe space restrictions. Such grammatical patterns as the BBI and the KDEC cover are found in general learner’s dictionaries, so that they may be left out unless more detailed semantic information about their use is provided.

It must be noted that ‘verb’ at the noun entry contains not only verbs, but also any kinds of phrases that take a headword noun as an object or a subject, and that can form the predicate of a sentence. For example, at the entry for rule, be bounded by, be fettered by, and be governed by are listed under the verb + rule section, and similarly, be applicable is listed together with apply and operate under the rule + verb section.

When a constituent of a collocation typically takes a certain grammatical form, it is shown in that form. For example, the collocation be teething at the baby + verb section at baby shows that the verb is typically used in the progressive in this collocation (viii). Another example can be taken from the adjective section in the entry for happy. It includes a list of adverbs that happy goes with typically in the negative: “far from, not altogether, not at all, not entirely, not exactly, not particularly, not too, not totally”. There is observed another way to indicate the grammatical context. At the entry for believe, among the adverbs seriously is listed with a gloss saying “only used with negatives”; however, no other example of this kind can be found by the writer.

6.2. ‘Word collocation’ and ‘category collocation’

The ties between the components of a collocation are lexical, as little cannot be used instead of small in small fortune though they are synonymous. Collocations of this type are called ‘word collocations’ in the OCDSE and equate to collocations in the general sense. The OCDSE recognizes and includes another type of collocations named ‘category collocations’.

There is another area of collocation that might be called ‘category collocation’, where a word can combine with any word from a readily definable set. This set may be quite large, but its members are predictable, because they are all words for nationalities, or measurements of time, for example. At the entry for walk, one of the groups of collocates is given as ‘three-minute, five minutes’, etc.: the ‘etc.’ is to indicate that any figure may be substituted for ‘three’ or ‘five’ in these expressions. (ix)

Some other examples of ‘measurements of time’ are 7-night, two-week etc. at the entry for holiday and three-year, two-year etc. at contract. This may be quite useful if users do not know the grammatical rule that attributive nouns are normally singular and proceeded by a hyphen (Quirk et al. 1985: 1333). Or some users may not know that they can simply say “a three-year contract” instead of “a contract valid for three years” or “a contract that runs [holds good] for three years”. Examples of the same kind are five-point, etc. at the entry for agenda, five-mile, etc. at run, and 200-piece, etc. at jigsaw. Ordinal numbers also form category collocations. For example, first, second, etc. is given under the adjective section of the noun entries for anniversary, act (sense 4), gear (sense 1), and round (sense 4). A set of adjectives such as weekly, monthly and yearly also belong to this group, and they are found at entries such as subscription (annual, monthly, yearly, etc.), tenancy (weekly, yearly, etc.), basis (daily, monthly, weekly, etc.), and check-up (annual, monthly, etc.). Below are shown some examples of entries that have ‘nationalities’ as collocates:

capital (British, Japanese, etc.), currency (European, Japanese, etc.), speaker (Japanese, Russian, etc.), ambassador (British, French, etc.), blood (Mediterranean, Spanish, etc.), silk (Chinese, Thai, etc.) passport (EU, Mexican, French, etc.), cheese (Dutch, Swiss, etc.), cooking (Chinese, Italian, etc.), ancestry (Celtic, French, etc.)

It is interesting to see national characteristics reflected in some entries. However, is it really necessary to include Dutch and Swiss at cheese and Chinese and Thai at silk? As for cooking, a phrase example “traditional English cooking” is provided for the adjective collocate traditional, and this may kill two birds with one stone. (Although “French regional cooking” is given as an example of Chinese, Italian, etc., this can be an example for regional and illustrate two collocations.)
7. Coverage

This section will look into some of the formal categories (cf. Section 6.1) and argue the utility of the kinds of collocations actually included from the perspective of foreign learners of English.

First, verb + noun collocations will be looked at. According to the semantic function they perform in a piece of writing, collocations can be divided into two types: collocations that convey particular ideas economically and precisely and those that add meanings or describe ideas more accurately or vividly (Hill et al. 2000: 93). For example, if a learner does not know the verb + noun collocation live up to expectations, it would be difficult to express the idea of it concisely and produce a sentence like The show failed to live up to our expectations. On the other hand, the adverb totally in I was totally exhausted at the time I advanced to the final merely intensifies the adjective exhausted, and is not an essential element to make the main point. Verb + noun collocations (as well as verb + adjective + noun collocations) can be considered to be "the single most important kind of collocation" (Hill et al. 2000: 116).

Second, adjective + noun collocations will be examined as they play both semantic functions mentioned above. Suppose a user is looking for an adjective to describe a change made which is only superficial and not important at the entry for change. The user may be unable to express his or her idea until he or she finds cosmetic. In this case, cosmetic change is a collocation that enables the user to name a concept. At the same time, it can be a collocation that enables us to express an idea more accurately. Suppose a user is not quite happy about describing a certain change as small or minor change. The user seeks for a more appropriate word and finds cosmetic, which may be better to mark his or her disapproval at the change.

Third, verb + adverb collocations and adverb + adjective collocations will be examined particularly in terms of intensification. Lorenz (1998: 53) explains the communicative effects intensification brings about as follows:

By amplifying and downtoning adjectival qualities, as in crucially important or hardly significant, we express assertion or caution, emphasis or doubt, and we take a committed or a non-committal stance towards the message in question. More than their mere denotation would suggest, intensifiers therefore convey speaker-stance, in some cases even to the point of creating a sense of identity and group membership.

Since correct use of intensifiers can be considered a step forward to natural English, it will be interesting to see how the OCDSE treats this issue.

Another type to be analyzed is verb + verb collocations, which are unique to the OCDSE, and they are not generally called 'collocations'. Lexical patterns treated here are quite different from ones in the other sections. What are included in this category seem to be auxiliary verbs, adjectives used predicatively followed by to-infinitive, verbs followed by to-infinitive, and anything that often or typically goes with the headword and that can form the predicate of a sentence as a whole. It will be discussed how helpful users would find this type of collocational information.

In order to throw light on character of the OCDSE, the analysis will be made in parts in comparison to two other dictionaries dealing (mainly) with collocations and one Japanese-English learner’s dictionary, though it is difficult to make a direct comparison since all the dictionaries that will be examined differ considerably in their editorial policies, compilation process, size etc. (See the table below.) It is claimed on the back cover that 150,000 collocations of 9,000 nouns, verbs and adjectives are provided with over 50,000 examples in context in the OCDSE. The BBI Dictionary of English Word Combinations (1997), a monolingual collocations dictionary for advanced learners of English, was compiled by three native speakers of English, with little or no access to corpus data, and makes a contrast to the OCDSE in that it relies largely on the compilers’ linguistic intuitions and theories. It is claimed on the front cover that the BBI contains 18,000 entries and 90,000 collocations. The number of entries is twice as large as that of the OCDSE, but the amount of collocational information per entry is much smaller. The Kenkyusha Dictionary of English Collocations (1995), a bilingual English-Japanese dictionary of collocations compiled on manually collected citations, will be also examined in order to see the difference or similarity between a monolingual...
collocations dictionary and a bilingual one. There is no mention of the number of headwords; however, according to my estimate based on sampling, the KDEC has about 19,295 headwords. The headword list consists mainly of nouns, verbs and adjectives, but includes a few adverbs, prepositions, and pronouns as well. All the collocations are basically shown in phrase or sentence examples in this dictionary, so that the number of examples amounts to as many as 380,000. Sometimes more than one example is given for one collocation, so that the number of collocations contained is likely to be smaller. However, the KDEC is the largest of all the dictionaries being discussed.

<table>
<thead>
<tr>
<th></th>
<th>OCDSE</th>
<th>BBI</th>
<th>KDEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>page size</td>
<td>233 mm × 155 mm</td>
<td>216 mm × 140 mm</td>
<td>247 mm × 170 mm</td>
</tr>
<tr>
<td>page (A to Z)</td>
<td>892</td>
<td>386</td>
<td>2,782</td>
</tr>
<tr>
<td>headwords</td>
<td>9,000</td>
<td>18,000</td>
<td>19,295</td>
</tr>
<tr>
<td>collocations</td>
<td>150,000</td>
<td>90,000</td>
<td>—</td>
</tr>
<tr>
<td>examples</td>
<td>50,000</td>
<td>—</td>
<td>380,000</td>
</tr>
</tbody>
</table>

In addition, the Kenkyusha Luminous Japanese-English Dictionary (2001), an encoding dictionary for Japanese learners of English, will be included in this comparison because it sees the importance of collocational information in writing and provides a number of boxed sections of collocations. For example, the entry for kiroku (=record) has a collocations box providing the following verb+noun collocations and adjective+noun collocations:

Lexical collocations at the entry for kiroku (= record) in the KLJED

verb + noun:
- 記録から漏れる (escape record)
- 記録する (makes record)
- 記録に留めておく (retain records)
- 記録を上回る (exceed the record)

adjective + noun:
- オリンピック記録 (the Olympic record)
- 簡潔な [詳細な] 記録 (a brief [detailed] record)
- 公式記録 (an official record)
- 高校記録 (a school record)

While the OCDSE and the BBI are targeted at learners of various nationalities, the KLJED tries to meet Japanese learners’ needs specifically. While the KDEC is rather for highly advanced learners of English such as academics and translators, the KLJED, a learner’s dictionary, may be considered to reflect the Japanese learners’ needs more.

7.1. Verb + noun collocations

A feature of the OCDSE as a corpus-based collocations dictionary becomes notable when its coverage of verb+noun collocations is compared to that of the BBI and the KLJED. Basically, the OCDSE aims to reflect typical or frequent use of word-combinations in a native-speaker corpus, which are what editors assume learners would need (cf. Section 3). As a result, the OCDSE covers fewer semantic categories in proportion to the number of collocates included, compared to non-corpus-based encoding dictionaries such as the BBI and the KLJED. For example, at the entry for rule, the OCDSE lists 35 verbs, which are grouped into ten semantic categories, under the first sense of ‘what you can or cannot do, say, etc.’ (see Table 1). For the corresponding sense, the BBI contains 20 verbs, and the KLJED, 14 verbs in the collocation box. In addition to the collocates contained in the OCDSE, the BBI lists set down and comply with, which should fall into group 1 and group 2 respectively, and four more collocates of three semantic groups: adopt, stretch, and rescind and revoke. The KLJED has prescribe and set up, which should belong to group 1, and respect and ignore can enter group 2 and 7, respectively. It has
Table 1: Comparison of verb collocates of rule

<table>
<thead>
<tr>
<th>OCDSE</th>
<th>BBI</th>
<th>KLJED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. draw up, establish, formulate, impose, issue, lay down, make, set out</td>
<td>draw up, establish, impose, lay down, set down</td>
<td>formulate, prescribe, set up</td>
</tr>
<tr>
<td>2. abide by, adhere to, follow, go by, observe, play by, stick to</td>
<td>abide by, obey, observe, comply with</td>
<td>respect</td>
</tr>
<tr>
<td>3. be in line with, conform to</td>
<td>be in breach of, break, disregard, fall foul of, flout, violate</td>
<td>be in breach of</td>
</tr>
<tr>
<td>4. apply, enforce</td>
<td>apply, enforce</td>
<td>apply, enforce</td>
</tr>
<tr>
<td>5. bend, relax</td>
<td>bend, relax</td>
<td>bend, relax</td>
</tr>
<tr>
<td>6. waive</td>
<td>waive</td>
<td>ignore</td>
</tr>
<tr>
<td>7. tighten up</td>
<td>tighten</td>
<td>tighten</td>
</tr>
<tr>
<td>8. be bound by, be fettered by, be governed by</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. interpret, understand</td>
<td>adopt, stretch</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td>remodel, rewrite, change</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>35</td>
<td>20</td>
</tr>
</tbody>
</table>

three more collocates, which could form another semantic category: remodel, rewrite and change. Although both the BBI and the KLJED contain much fewer verb collocates overall, the BBI covers nine semantic categories, and the KLJED, eight. Since the main function of verb + noun collocations is to name a concept that a user has in mind, the coverage of broader semantic categories may be given priority over that of collocations that have high frequencies in the corpus from the same semantic group.

Also, collocations included in the OCDSE may differ in some cases considerably from those which Japanese lexicographers predict that users would find useful. It is interesting to see that only four out of the eleven verb collocates of the noun delay included in the OCDSE overlap with those in the collocation box in the KLJED. The verbs that appear in the both dictionaries are asterisked.

Table 2: Comparison of verb collocates of delay between the OCDSE and the KLJED

<table>
<thead>
<tr>
<th>OCDSE</th>
<th>KLJED</th>
</tr>
</thead>
<tbody>
<tr>
<td>be subject to, be plagued by, experience, face, suffer*, cause*, lead to, avoid*, prevent, reduce*, apologize for</td>
<td>suffer*, cause*, avoid*, reduce*, brook no —, create, produce, eliminate, give rise to, increase</td>
</tr>
</tbody>
</table>

A similar difference can be observed when the verb collocates of the noun compliment are compared to those in the BBI, the KLJED, and the KDEC. In Table 3 below, the collocates are semantically grouped together, on the basis of the categories in the OCDSE and the BBI. A glance at the table tells that the coverage of the OCDSE is different from the other three. Only three out of ten semantic categories are covered by all the four dictionaries, and the three collocations pay somebody a compliment, accept a compliment, and return a compliment are included in the all of them. One collocation, mean something as a compliment, is unique to the OCDSE. Noticeably, the BBI and the KLJED have a strikingly similar coverage. Since they are both not corpus-based, it may be said that editors' view about learners' needs are reflected more directly in the types of collocations included. It is interesting that a collocations dictionary compiled by native speakers and a Japanese-English encoding dictionary compiled by non-native speakers agreed on almost the same inclusion. The KDEC, the biggest of all, provides the most collocates, and is similar to the BBI and the KLJED in its coverage. Shower, angle for, fish for a compliment, and bandy compliments are included in the all dictionaries except the OCDSE, which may carry some implications for inclusion.
Table 3: Comparison of verb collocates of *compliment*

<table>
<thead>
<tr>
<th></th>
<th>OCDSE</th>
<th>BBI</th>
<th>KLJED</th>
<th>KDEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pay sb</td>
<td>pay smb</td>
<td>pay</td>
<td>pay</td>
</tr>
<tr>
<td></td>
<td></td>
<td>make</td>
<td>make</td>
<td>offer</td>
</tr>
<tr>
<td>2</td>
<td>shower</td>
<td>shower</td>
<td>shower</td>
<td>shower</td>
</tr>
<tr>
<td></td>
<td>lavish</td>
<td>lavish</td>
<td>lavish</td>
<td>lavish</td>
</tr>
<tr>
<td></td>
<td>shower</td>
<td>heap</td>
<td>heap</td>
<td>offer</td>
</tr>
<tr>
<td></td>
<td>shower</td>
<td>shower</td>
<td>shower</td>
<td>offer</td>
</tr>
<tr>
<td></td>
<td>shower</td>
<td>shower</td>
<td>shower</td>
<td>offer</td>
</tr>
<tr>
<td></td>
<td>shower</td>
<td>shower</td>
<td>shower</td>
<td>offer</td>
</tr>
<tr>
<td></td>
<td>shower</td>
<td>shower</td>
<td>shower</td>
<td>offer</td>
</tr>
<tr>
<td></td>
<td>shower</td>
<td>shower</td>
<td>shower</td>
<td>offer</td>
</tr>
<tr>
<td>3</td>
<td>mean sth as</td>
<td>get</td>
<td>receive</td>
<td>get</td>
</tr>
<tr>
<td></td>
<td></td>
<td>receive</td>
<td>receive</td>
<td>get</td>
</tr>
<tr>
<td>4</td>
<td>accept</td>
<td>accept</td>
<td>accept</td>
<td>accept</td>
</tr>
<tr>
<td></td>
<td>acknowledge</td>
<td>accept</td>
<td>receive</td>
<td>accept</td>
</tr>
<tr>
<td>5</td>
<td>regard sth as</td>
<td>take smt. as</td>
<td>take as</td>
<td>echo</td>
</tr>
<tr>
<td></td>
<td>take sth as</td>
<td>take smt. as</td>
<td>take as</td>
<td>echo</td>
</tr>
<tr>
<td>6</td>
<td>repay</td>
<td>return</td>
<td>return</td>
<td>return</td>
</tr>
<tr>
<td></td>
<td>return</td>
<td>return</td>
<td>return</td>
<td>return</td>
</tr>
<tr>
<td>7</td>
<td>repay</td>
<td>return</td>
<td>return</td>
<td>return</td>
</tr>
<tr>
<td>8</td>
<td>angle for</td>
<td>angle for</td>
<td>angle for</td>
<td>angle for</td>
</tr>
<tr>
<td></td>
<td>fish for</td>
<td>fish for</td>
<td>fish for</td>
<td>fish for</td>
</tr>
<tr>
<td>9</td>
<td>bandy s</td>
<td>bandy s</td>
<td>bandy s</td>
<td>bandy s</td>
</tr>
<tr>
<td>10</td>
<td>shrug off</td>
<td>lend off</td>
<td>shrug off</td>
<td>lend off</td>
</tr>
<tr>
<td>total</td>
<td>10</td>
<td>9</td>
<td>11</td>
<td>18</td>
</tr>
</tbody>
</table>

7.2. Adjective + noun collocations

Adjectives in adjective + noun collocations can be divided into two types: quality-modifying adjectives and category-naming adjectives. For example, the entry for the noun *speech* lists the following 30 adjectives for the sense of ‘formal talk’:

brief, little, short | interminable, long, long-winded, rambling | keynote, major | eloquent, excellent, good | emotional, impassioned, rousing | striking | boring | impromptu | public | televised | political | opening | closing | acceptance, after-dinner, Budget, campaign, con-

From *brief* to *impromptu* are adjectives that modify the content of a speech, and from *public* to *inaugural* are adjectives that describes the type of a speech. These two kinds of adjective collocates serve quite a different function. In the Introduction, the *OCDSE* explains that “A student who chooses the best collocations will express himself much more clearly and be able to convey not just a general meaning, but something quite precise” (vii) by comparing the grammatically correct two sentences:

This is a good book and contains a lot of interesting details.

This is a fascinating book and contains a wealth of historical detail.

The use of quality-modifying adjectives promotes preciseness or vividness of expression. On the other hand, category-naming adjectives may greatly contribute to succinctness. For example, if learners do not know the collocation *inaugural speech*, then it would be difficult to describe succinctly the first speech made by a new president at the ceremony.

As for the coverage, there seems to be neither clear difference nor strong similarity between the *OCDSE*, the *BBI*, and the *KDEC*, when the entries for *speech* (in the sense of “formal talk”) are compared (see Table 4). The *KLJED* has the following adjectives at the entry for *enzetsu* (speech) and most of them are more or less the same as the ones in the table: *welcome* [welcoming], *stirring* [moving], *keynote* *inaugural acceptance, political, campaign, agitative, presidential, funeral, long and boring, poor, short* [brief], pointless. Unlike the case of verb + noun collocations, it seems that generally, the more collocates are covered, the more semantic categories are covered in the case of adjective + noun collocations in the *OCDSE*.

7.3. Adverb + adjective collocations

The *OCDSE* recognizes importance of intensifying adverbs and gives a good coverage of them in general. Here are some examples:

impressive, decidedly, enormously, extremely, highly, hugely, immensely, mightily, most, particularly, really, tremendously, truly, very

rich

enormously, extremely, fabulously, filthy (informal), im-
Although its full treatment of intensifying adverbs is a welcome policy, the OCDSE also has some adverbs whose inclusion needs to be questioned. First, as all the above examples show, the OCDSE includes very as a possible choice at a number of entries. Some entries such as cloudy, fixed and handy give only very as intensifiers. Probably, it is more instructive and space-saving to indicate the unacceptability of use of very when it cannot be used, since learners are not likely to be aware of the fact that even an apparently all-purpose intensifying adjective very is by no means universally applicable as one cannot say, for example, *very awake, *very asleep, or *very apart (Sinclair et al. 1992: 758). Second, taboo words such as bloody and fucking may not be suitable for inclusion, when the dictionary claims that the type of texts its target users are likely to produce is “what might be called ‘moderately formal language’ — the language of essay and report writing, and formal letters — treating all subjects — business, science, history, sport, etc . . . . at the level of the educated non-specialist” (viii–ix). In addition, bloody and fucking go with almost any word. Third, a set of adverbs of ‘degree’ are often given at entries for adjectives that describe feelings; however, it is doubtful whether they merit inclusion in a collocations dictionary as they combine with quite a few adjectives without lexical restrictions.

Although its full treatment of intensifying adverbs is a welcome policy, the OCDSE also has some adverbs whose inclusion needs to be questioned. First, as all the above examples show, the OCDSE includes very as a possible choice at a number of entries. Some entries such as cloudy, fixed and handy give only very as intensifiers. Probably, it is more instructive and space-saving to indicate the unacceptability of use of very when it cannot be used, since learners are not likely to be aware of the fact that even an apparently all-purpose intensifying adjective very is by no means universally applicable as one cannot say, for example, *very awake, *very asleep, or *very apart (Sinclair et al. 1992: 758). Second, taboo words such as bloody and fucking may not be suitable for inclusion, when the dictionary claims that the type of texts its target users are likely to produce is “what might be called ‘moderately formal language’ — the language of essay and report writing, and formal letters — treating all subjects — business, science, history, sport, etc . . . . at the level of the educated non-specialist” (viii–ix). In addition, bloody and fucking go with almost any word. Third, a set of adverbs of ‘degree’ are often given at entries for adjectives that describe feelings; however, it is doubtful whether they merit inclusion in a collocations dictionary as they combine with quite a few adjectives without lexical restrictions.

7.4. Verb + verb collocations
The purpose of the inclusion of ‘verb + verb’ collocations, which is unique to the OCDSE, may be to enable users to produce natural English, that is a step forward from producing grammatically correct English or a lexically acceptable combination of words. However, there also seem to be some problems about selection of the items. The followings are examples of ‘verb + verb collocations’ randomly collected:
VERB + ACCEPT (sense 2) be happy to, be prepared to, be ready to, be willing to, be reluctant to, be unwilling to, be unable to, refuse to, can/cannot, could (not), will/won't, would (not), be forced to be

VERB + ADMIT be forced to, have to, must, refuse to, be honest enough to, be prepared to, be the first to, be willing to, dare (to), have the courage to, be ashamed to, be embarrassed to, be loath to, be reluctant to, be unwilling to, hate to, not care to

VERB + AGREE (sense 1) cannot, be inclined to

VERB + OPPOSE vote to, continue to

VERB + ORGANIZE seek to, try to, help (to), be able to, manage to, fail to, learn (how) to, be difficult to, be easy to, be possible to

VERB + REFUSE can't/couldn't, can/could hardly, be entitled to, have the right to, be/seem churlish to

VERB + THANK must, want to, would like to

Certainly, the phrases listed above form a recurrent phrase with the headword; however, in some cases what connects a sequence of words that commonly co-occur is more grammatical than lexical. For example, be difficult to, be easy to, and be possible to at organize seem to be syntactic strings, though they may be found typical in the corpus data, and the inclusion of these phrases in a collocations dictionary is therefore questionable. Some items seem to have a pragmatic function when they are used at the sentence level. For example, want to at thank forms a spoken phrase “I want to thank...” used when you express your gratitude. At the entry for want in the Macmillan English Dictionary (2002) “I want to say/thank/tell etc.” is given as a spoken phrase and is explained as “used for introducing something that you are about to say, especially at the start of a speech” with an example “I want to thank you all for being here.”. Another example is have to and must at the entry for admit. The Longman Advanced American Dictionary (2000) explains “I have to admit” is “used when you are admitting something you are embarrassed about” with an example: “I have to admit I was a little drunk.”. However, inclusion of these phrases sometimes does not seem consistent. If be willing to is included at accept, why not at help? If be ashamed to is included at admit, then why not at say? If must is included under thank, then why cannot, which makes an idiomatic expression “I can’t thank you enough for...”, is not?

Inclusion of this type of phrases is beneficial to learners; however, it is quite doubtful whether users can actually put them into the typical context they are used in as they are not always accompanied by examples. At the entry for agree, phrases with a pragmatic function such as couldn’t agree more, I quite agree and I’m sure you will agree are given under phrases with examples: I couldn’t agree more with what has just been said. And I’m sure you’ll all agree that this issue is vitally important to the success of the company. This approach seems better since these phrases are probably hard to be understood without any contextualization.

7.5. Conclusion
Although only four types of collocations have been examined, there seems to be a little bit too much emphasis placed on the corpus data in regard to coverage, as some of the examples cited in the above sections illustrate. Lea and Runcie (2002: 824) say that “What learners of English really need to know are the collocations that native speakers actually use, in the course of writing and speaking on any subject but collocation; not necessarily what ‘springs to mind’ when a native-speaker is asked to name a collocation.” Acquisition of collocations frequently used by native speakers will certainly contribute to fluency. However, it does not mean that a good learner’s collocations dictionary has only to represent accurately frequent, typical word combinations in a native-speaker corpus. Some differences observed between the OCDSE and the other three dictionaries may also imply that users’ needs cannot be covered simply by providing frequent collocates retrieved from a native-speaker corpus, which is designed to be a representative sample of the current English. Although the OCDSE does provide good models with learners, taking the foreign learner’s viewpoint may lead to improvement.

8. Suggestions for improvements
In this last section, I would like to make some suggestions on how to make
the dictionary an even more valuable tool. First, it is regrettable that the
OCDSE provides no cross-references between related or synonymous
headwords. With cross-references, more information would be accessible
to users. For example, if the entry for movie is cross-referred to that for
film, users can simply gain much more information. In addition, a cross-
reference from a general word to a specific word may help users to make a
more appropriate choice. Suppose that a user is trying to lexicalize a
feeling of great anger and looks up the entry for angry. The user may be
happy finding collocations such as bitterly angry and extremely angry.
However, a cross-reference from angry to furious, for example, proposes
another solution to the user, that is a choice of a different lexical item (as
a starting point) to express the idea.

Second, information about incorrect use may be appreciated because
collocations are by nature open-ended and no list could ever be exhaustive
or definitive. In other words, all users find is (lists of) what is possible or
correct, and here arises a problem. When a word combination a user has in
mind is not listed in a dictionary, it may not be in because it is deviant. It
may not be included because it is a free combination or on the borderline,
but still is fully acceptable. There is no way for users to check whether or
not the combination is acceptable. The Longman Essential Activator (1997),
an encoding dictionary, tries to prevent learners from making mistakes by
giving notes called ‘Essential Help Boxes’ based on the error analyses of
the Longman Learner’s Corpus. For example, at the concept crime, a help
box reads: “Don’t say ‘he made a crime’ or ‘he did a crime’. Say he committed a crime.” At the concept amount/number the note is given
that “Don’t say ‘a big number’ Say a large number.” Taishukan’s
Genius Japanese-English Dictionary, second edition (2003), a Japanese-
English learner’s dictionary for production, takes a similar approach. It
marks “less-typical” collocations with a triangle (°) and “non-typical”
collocations with a cross sign (x) in order to show the degree of acceptabil-
ity (x). For example, the entry for kibishit (severe, strict, rigid, etc.) has an
example to show less-typical use in contrast to typical ones: The doctor ordered her to go on a rigid [strict, °severe] diet. At the entry for katai (hard,
firm, rigid, etc.), a set of adjective collocates to go with the noun cover are
shown: a book with a hard [stiff, tough, °rigid, °firm, °solid] cover. Some
examples of incorrect collocations are: make [°do] two mistakes in arith-
metic at machigau (mistake), enter [°join] school, join [°enter] the line at
hairu (enter, join), and She was very (much) °badly, °seriously hurt to hear
him say so. at kibun (mood, feeling, etc.). These usage notes are made
based on native speakers’ intuitions, corpus data, and Japanese lexicographers’ insights into learners’ needs (x).

Third, more semantic information about synonymous collocates may be
necessary in order to prevent users from making dictionary-oriented
mistakes. The presentation of synonymous collocates in the same list may
lead users to think that all the collocates are interchangeable, when they
are not at all. For example, at section 1 of the entry for change, seven
adjectives meaning ‘small in degree and/or of little importance’ (cosmetic,
marginal, minimal, minor, slight, small, subtle) form one subsection. How-
ever, while small is a neutral word, cosmetic has a negative prosody.
Although it is suggested in ‘Introduction’ to consult a general learner’s
dictionary such as the Oxford Advanced Lerner’s Dictionary, in order to
make the best choice, a general EFL dictionary sometimes fails to give
distinctive semantic features (cf. Komuro 2003: 134–35) to near-syn-
onyms. Moreover, it is more desirable for a learner’s dictionary to be self-
contained.

The last suggestion to make concerns the entry structure. Although the
OCDSE and other collocations dictionaries usually group collocates to-
gether and present them according to their parts of speech, the themati-
ically structured entry might be able to help users more with their
point out, it cannot be assumed that learners know from the outset the
precise grammatical form that the lexical realization of their idea might
take. The Longman Essential Activator presents words and phrases that
share the same idea irrespective of grammatical categories with a heading
to name the category, which may be applied to the entry structure of a
collocations dictionary. Below is a proposed entry for the semantic unit
refuse based on the data in the OCDSE, reorganized to give priority to
meaning over form.
REFUSAL noun, REFUSE verb

1. to refuse something in a very clear and definite way
   refuse flatly/point-blank/categorically (= refuse in a firm and sometimes impolite way)
   He flatly refused to discuss the matter.
   refuse point-blank to co-operate.
   give a flat/point-blank/complete refusal
   refuse pointedly (= refuse clearly in a way that shows you are annoyed)
2. to be refused
   be met with by/with a refusal
   Her appeals for funds to support the cause were met with blank refusal.

The entry shows the users that there are (at least) two different grammatical structures to express the idea of refusing something in a very clear way. For the long-term benefit of the learners, it might be more effective to show lexical bonds that lie beyond parts of speech in order to develop users' collocational competence as Stubbs (2002: 30) states that "the collocation is between semantic units, irrespective of grammatical category".

9. Concluding remarks
It can be said that the significance of collocations has been increasingly recognized in the Oxford tradition. The **Oxford Dictionary of Current Idiomatic English** (Volume 1: Verbs with Prepositions & Particles, 1975; Volume 2: Phrase, Clause & Sentence Idioms, 1983) is "the first large-scale phraseological dictionary of English to be compiled by native speakers" (Cowie 1998: 220). 'General Introduction' shows its foresight to see pedagogical importance of acquisition of collocations, as it begins with "Familiarity with a wide range of idiomatic expressions, and the ability to use them appropriately in context, are among the distinguishing marks of a native-like command of English" (vi). A critical appraisal of the **Oxford Advanced Learner's Dictionary of Current English**, the 4th edition (1989) by Takahashi et al. (1992) points out its good coverage and treatment of collocations. Its fifth and sixth editions published in 1995 and in 2000 respectively, contain a language study page 'Collocation' consisting of three subsections: 'What is collocations?', 'Types of collocations' and 'Collocation in this dictionary'. A good collocations dictionary was therefore expected to come out from Oxford, and the **OCDSE** lives up to our expectations. The **OCDSE** is probably of almost maximum size as a portable collocations dictionary, and its presentation of information is typographically neat. This dictionary should go in the list of dictionaries recommended to learners of English. In order to make this dictionary more accessible for less advanced learners, bilingualized versions would be widely welcomed.

**NOTES**

1) I would like to express my deep sense of gratitude to Emeritus Prof. Keisuke Nakao for his suggestions, comments and encouragement at every stage of this study. I am also truly grateful to Mr. Michael Rundell for his comments and great help with the final draft.
2) This dictionary was compiled on the basis of **Selected English Collocations** (1982) and its companion volume, **English Adverbial Collocations** (1991).
3) Although it is not mentioned which edition was used, it may be assumed that the second edition published in 1993 was made use of for the compilation.
4) There is no mention of which edition was used, but it may be assumed that the latest version at the time, that is the 6th edition published in 2000 was used.
5) Homonyms are dealt with under the same entry. For example, the noun entry bank has four sections: “1 for money”, “2 by a river/riverbank”, “3 area of sloping ground”, and “4 mass of cloud, etc./row of machines, etc.”.
6) I selected ten pages from every hundredth page from page 1 onwards as sample materials and counted headwords, which are printed in bold. The table below shows the samples pages and the number of headwords in each section.

<table>
<thead>
<tr>
<th>page</th>
<th>from</th>
<th>to</th>
<th>no. of headwords</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-10</td>
<td>abacus</td>
<td>abuse</td>
<td>74</td>
</tr>
<tr>
<td>200-210</td>
<td>babble</td>
<td>bank</td>
<td>80</td>
</tr>
<tr>
<td>400-410</td>
<td>charge</td>
<td>child</td>
<td>73</td>
</tr>
<tr>
<td>600-610</td>
<td>cross</td>
<td>curious</td>
<td>85</td>
</tr>
<tr>
<td>800-810</td>
<td>dress</td>
<td>duckling</td>
<td>53</td>
</tr>
<tr>
<td>1000-1010</td>
<td>field</td>
<td>finger</td>
<td>57</td>
</tr>
<tr>
<td>1200-1210</td>
<td>help</td>
<td>hiss</td>
<td>87</td>
</tr>
<tr>
<td>1400-1410</td>
<td>launch</td>
<td>leaning</td>
<td>45</td>
</tr>
<tr>
<td>1600-1610</td>
<td>money maker</td>
<td>morning</td>
<td>52</td>
</tr>
<tr>
<td>1800-1810</td>
<td>pilgrimage</td>
<td>place</td>
<td>62</td>
</tr>
<tr>
<td>2000-2010</td>
<td>queen</td>
<td>radio</td>
<td>68</td>
</tr>
<tr>
<td>2200-2210</td>
<td>scull</td>
<td>secret</td>
<td>54</td>
</tr>
</tbody>
</table>
The average number of headwords per page is 6.935714, and the estimate number of the total headwords is about 19,295.

7) Quoted from the entry for refuse of the Macmillan English Dictionary (2002).

REFERENCES