A Corpus-Based Approach to the Vocabulary Teaching and Learning—Implications From the Corpus-Based Collocational Research on Two Pairs of Synonyms

CHEN Yan-hui
Shanghai Institute of Technology, Shanghai, China

Languages embody a lot of words that are considered as synonyms, and people just take it for granted that such words are identical in meaning without any discrimination. However, a corpus-based approach to the study of the collocational behavior of the two frequently-used pairs of synonyms (selection and option, ill and sick) reveals significant discrepancies in the use of these two pairs of synonyms by Chinese English learners and native speakers. According to the analysis, the major problems lie in the current ways of vocabulary teaching and learning. This paper aims to highlight the important role of the corpus-based collocational research in English vocabulary teaching and learning. In the end, some suggestions concerning vocabulary teaching and learning are put forward on the basis of corpus-based research.

Keywords: corpus, collocation, synonym, vocabulary teaching

Introduction

With the development of the corpus linguistics, a corpus-based approach to the English teaching and learning is becoming more and more important in recent years. “A corpus is a collection of naturally-occurring language text, chosen to characterize a state or variety of a language” (Sinclair, 1999, p. 26). It has obviously had an impact in a number of fields in the language sciences. In contrast to Chomskyan approaches to language, corpus-based descriptions are based on true and authentic linguistic performance as the source of evidence for theories of language, and so far have largely focused on particular languages rather than universals of language. Therefore, a corpus-based approach to the language teaching and learning is helpful and meaningful in many aspects in real practice. For instance, English learners may find it easy to learn and memorize the meaning of a word with the help of dictionaries and thesauruses. However, it might not be that easy to distinguish synonyms and use them accurately. Most dictionaries and thesauruses only provide a list of synonyms without clear distinction between them. Teachers also give students a lot of synonyms without clear-cut distinction between them. Thus, students may get confused and have problems in using these synonyms. However, through analysis of frequency distribution of the synonyms, some very distinctive preferences in their usage can be demonstrated.

Adopting the corpus-based approach, the present study analyzes the collocational behavior of the two
frequently-used pairs of synonyms (*selection* and *option*, *ill* and *sick*). The significant discrepancies in the use of these two pairs of synonyms by Chinese English learners and native speakers will be revealed, and the important role of the corpus-based collocational research in English vocabulary teaching and learning will be illustrated as well. The pedagogical implication of the study is that emphasis and concentration on the collocations and usage of synonyms in light of their actual context and collocational restrictions will improve the learners’ vocabulary learning efficiency and enhance their writing ability.

**Materials for Research and Ways of Researching**

Adopting the corpus-based approach, this paper tries to provide a teaching module that might be beneficial to the teaching and learning of vocabulary, especially synonyms. In order to have a thorough analysis of the difference in using synonyms between English learners and native speakers, two pairs of synonyms are carefully selected and thoroughly analyzed. All the examples and statistics used in this paper are chosen from British National Corpus (BNC) and Collins Birmingham University International Language Database (COBUILD), a collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of current British English, both spoken and written.

As John Sinclair (1999) suggested, “learners would do well to learn the common words of the language very thoroughly, because they carry the main patterns of the language. The patterns have to be rather precisely described in order to avoid confusions” (p. 75). Therefore, two very common pairs of synonyms (*option* and *selection*, *ill* and *sick*, the former are two nouns, the latter are two adjectives) are chosen, which are usually told with the same usage in daily life. In the following, the ways of analyzing and learning pairs of synonyms will be illustrated. The analysis of synonyms will be made from three aspects, say, sense, grammar, and collocation to investigate the different usages between English learners and native speakers. At the same time, a comparison of different use of these two pairs of synonyms is drawn between Chinese learners and native English speakers. Also, the objective of this research is to suggest the pedagogical implications of the findings.

**A Corpus-Based Analysis of Two Pairs of Synonyms**

Both words have a meaning of the act of choosing. Look at the following examples from *Longman Dictionary of Contemporary English* (2004).

Example (1) The most important thing is your firm knowledge that the *selection* is within your present range.

Example (2) As for replacement fuels, many people do not like to contemplate the nuclear *option*.

Firstly, the senses of *selection* and *option* will be examined. The principal sense of these two words is with no difference, which is the act of choosing and selecting, but *option* has a strong association with economic discipline. It means the exclusive right, usually obtained for a fee, to buy or sell something within a specified time at a set price. For example:

Example (3) Each company has granted the other a 19% stock *option*.

Example (4) Consequently share *option* schemes are a thousand better than profit-sharing schemes.

This sense cannot be ignored, because it occurs with a high frequency in the corpus. Actually, “share option” and “stock option” rank first and fourth in the list of “noun + option” according to frequency (see Appendix A).
Several different lists of words appeared before *option* and *selection* according to frequency are provided, but unexpectedly, when “stock option” is mentioned, very few English learners heard of this phrase, let alone knew the exact meaning of this phrase. That is what can be done with a corpus, to provide most widely used language materials to English learners.

Secondly, the subject of *selection* and *option* will be concentrated. Compared with *option*, *selection* tends to associate with words with good sense. According to Appendix A, the most frequently appeared adjectives used before *selection* are words with good sense, such as wide, good, fine, large, careful, wonderful, extensive, excellent, interesting, best (see Appendix A).

In comparison to the *selection*, *option* is more used with neutral meaning. Take some adjectives as example, such as military, only, soft, viable, nuclear, cheaper, practical, likely, negative, and real (see Appendix A).

Thirdly, the collection of *selection* and *option* is illustrated. According to Appendix A, it is found that *selection* and *option* differ so much in collection. For instance, people can say “make selection” but not “make option”, because according to the statistics provided by corpus, “make” is the most frequently used verb with *selection*, followed by carry, receive, include, show, find, see, choose, print, serve, and so on; while “offer” ranks the highest in use with *option*, followed by retain, exercise, consider, get, keep, take, demand, include, want, and pursue.

A dramatic difference found from the COBUILD is that *option* usually associates with “share”, while *selection* is used with “stock”, as is shown in Table 1 from Collins Word Banks Online English Corpus.

| The Number of Samples of “Share Option” and “Stock Selection” in COBUILD |
|-------------------------------|------------|-----------|-------------|
| Collocate | Corpus Freq. | Joint Freq. | Significance |
| Option     | Share       | 5,442     | 44          | 6,352,303    |
| Selection  | Stock       | 3,593     | 14          | 3,340,240    |

Last but not least, the grammatical restriction on the *selection* and *option* is also different. According to Appendix A, the research reveals that *selection* and *option* are used in different forms. In order to show this difference clearly, some examples and statistics from BNC are employed to illustrate this point. The single form of these two words appear nearly the same times in BNC, with *option* 5,392 times and *selection* 5,993 times. For example:

Example (5) She knew she would not cheat and take that *option*, and she could think of no other.

Example (6) All that is recommended is that experienced, practicing teachers be increasingly involved in the *selection* process.

However, there is a sharp difference between their plural forms: *Options* appears 3,749 times in the corpus, while *selections* only 222 times. Thus, it can be concluded that *selection* is more likely to be used in singular form rather than plural form. Also, it is found that the propositions used after each phrase is so different. According to the examples and data selected from BNC, it reveals that “a selection” has a very strong link with “of”, which can be regarded as an idiomatic phrase, while “an option” does not have such strong link with prepositions.

From the above analysis, it is revealed that there is much should be covered when teaching students these two words, though they have the same principal meaning as the act of choosing. *Option* has a strong association
A CORPUS-BASED APPROACH TO THE VOCABULARY TEACHING AND LEARNING

with economic discipline. It means the exclusive right, usually obtained for a fee, to buy or sell something within a specified time at a set price. Phrases like “share option” and “stock option” appear very frequently in actual use. Selection tends to associate with words with good sense; while option is more often used with neutral meaning. Selection is more likely to be used in singular form rather than plural form. The proposition used after “a selection” and “an option” is very different: The former has a strong link with “of”, while the latter does not have such strong link with proposition. As for collocation, the two words differ significantly. They collocate with different words, including nouns, adjectives, and verbs.

When Chinese learners learn the two words, the above difference and idiomatic usage are usually overlooked. Therefore, great importance should be attached to such differences when teaching and learning synonyms. In order to make the point more convincing, just have a look at the following two words, ill and sick.

Firstly, the senses of ill and sick are examined. According to Longman Dictionary of Contemporary English (2004), these two words are used as adjectives in most cases, with the meaning of “not feeling well, having some kind of disease or a problem with one’s health”. For example:

Example (7) All these diets are making you ill.
Example (8) I was off sick for four days with the flu.

But sick has another meaning which means to bring up food through your mouth from your stomach. For example:

Example (9) You’ll be sick if you eat any more of that chocolate!

Ill also has minor sense besides the principal one, that is, the sense of being bad, harmful, and cruel. From the 100 concordances of ill, “ill health” appeared seven times, “ill treatment” three times (see Appendix B).

Secondly, what kind of subjects are used with these two words will be discussed. According to the concordances of ill and sick from corpus, sick and ill are usually used to describe people in bad condition. Though their collocations normally do not have preference of whether they are good or bad in sense, once they are modified by these adjectives, their collocations appear to be negative.

Then, the collocation of ill and sick will be discussed. According to the corpus, ill and sick are used as adjectives in most cases. As for ill, 98 out of 100 are used as adjectives; as for sick, 96 out of 100 are used as adjectives. However, the frequency of noun form of ill and sick is too low, only two and four out of 100. It is found that most of sick and ill are used as predicative, but still some are used as attribute. It is also revealed that the frequency of “sick + noun” is much higher than “ill + noun”. “Sick + noun” appeared 31 times; “ill + noun” appeared 18 times in the corpus. Seeing thoroughly, it can be found that there are still differences in sense. When sick is used as attribute, in most cases, it has the meaning of having some kind of disease and usually collocates with living things such as animals, human beings. Different from this, ill tends to appear more likely with the sense of bad, harmful, and cruel in sense. For instance, “ill health” appeared seven times in the corpus. It is worthwhile to mention this collocation, when describing a person’s health status; Chinese learners usually use “bad health” or “poor health”. However, there is no such collocation from the material. Surprisingly enough, “ill health” appears seven times in the material. In order to use more convincing data to illustrate this, the word ill is searched in BNC, finding “bad health” appears 20 times and “poor health” 121 times, however, “ill health” appears 320 times in BNC (see Figure 1). Apparently, ill strongly associates with health, which Chinese learners often neglect this point.
Thus, when teaching or learning the word *ill*, much attention should be paid to this collocation. Moreover, *ill* is usually used with adverbs like seriously, terminally, desperately, severely, or mentally when it has the meaning of having some kind of disease. While, *sick* is used with adverbs like chronically, physically, desperately, continually, really, personally, or mentally. But adverbs like desperately, mentally appeared more frequently with *ill* than *sick*. The preposition used after *ill* and *sick* is also different. The preposition appeared after *ill* like in phrase “speak ill of”, “ill at ease” or “ill + with”, “ill + for”; the preposition appeared after *sick* like in phrase “be sick of” or “sick + to” (see Appendix B). In conclusion, though these two words are synonyms, they have different usages and collocations.

In conclusion, though these two words are synonyms, they have different usages and collocations.

**Pedagogical Implication of the Corpus-Based Analysis**

From the thorough analysis of the two pairs of synonyms, the current ways of teaching vocabulary should be reconsidered and reexamined carefully, especially synonyms. Memorizing a list of synonyms without distinction is not a correct way to learn English, because these seemingly synonymous words typically used in very different ways. Each is particular and has its uses and idiosyncratic environment. There are some aspects to which more attention should be paid in English vocabulary teaching.

Since synonyms usually embody exactly the same meaning. However, among one entry, one word may have a sense, besides principal senses, that differs from other, which should be taken into account when teaching or learning synonyms, because that is also evident to tell synonyms apart.

The second point is that the subject of a word should be emphasized, because the most distinctive features of a word can be demonstrated through its subject. As is known to all, some words tend to associate with words with good sense, others unpleasant sense. For instance, the usual subjects of “set in” are rot, decay, malaise, despair, ill-will, decadence, impoverishment, infection, prejudice, vicious, rigor mortis, bitterness, anticlimax, disillusion, slump, mannerism, and anarchy (Sinclair, 1999). Not one of them is conventionally desirable or attractive.

The third point concentrates on the fact that in teaching and learning vocabulary is word-form and sentence pattern. Some words are sensitive to the passage of time, while others are not. Many synonyms may tend to appear in different word forms: Some are more constantly used in the past; some are not; some may not be used in speculation about the future, or in statements about the present. With the grasp of this information, students are
more likely to choose more suitable word. The features of grammar can be distinguishing criteria of synonyms. Some words, especially verbs and conjunctions tend to have very different grammatical patterns. There is much to consider in this aspect when learning or teaching a word, which helps to present the facts of the language in a much more precise way, because words can be acquired in typical contexts.

Last but not least, a corpus-based approach should be adopted in vocabulary teaching and learning. The collocation of a word shows that words do have a preference with certain words. Many synonyms have big differences in their collocations, which can be easily proved by statistics of a corpus. If teachers and students know the collocation of words, it will not be that hard to distinguish synonyms. Students will be able to produce with confidence much more idiomatic English, with less processing effort.

The above four aspects are very essential in teaching and learning the vocabulary. These observations and revelations should be presented to English learners in classroom teaching. Some of the findings of corpus-based investigation are seldom mentioned in current teaching materials. What is more, it is better if this method can be adopted into lexicography since many teachers and students refer to dictionaries of synonyms as their guidance in English learning. Through the thorough analysis of the two pairs of synonyms, an overview picture of corpus-based vocabulary teaching and learning is revealed clearly.

**Conclusion**

Adopting the corpus-based approach, this paper analyzes the collocational behavior of the two frequently-used pairs of synonyms (*selection* and *option*, *ill* and *sick*). The analysis reveals significant discrepancies in the use of these two pairs of synonyms by Chinese English learners and native speakers. According to the above analysis, the major problems lie in the current ways of vocabulary teaching and learning. The present vocabulary teaching still has room for improvement. Therefore, based on the corpus-based approach, a teaching module that might be beneficial to the teaching and learning of vocabulary, especially synonyms, is provided. The pedagogical implication of the study is that emphasis and concentration on the collocations and usage of synonyms in light of their actual context and collocational restrictions will improve the learners’ vocabulary learning efficiency and enhance their writing ability. Some suggestions are also put forward on the basis of the study. For instance, teachers should adopt a corpus-based approach to the vocabulary teaching. Because the traditional vocabulary teaching seldom uses authentic and idiomatic language material, however, corpus provides English learners with an enormous warehouse of genuine language materials, which is extremely essential to both language teaching and learning. Moreover, the vocabulary teaching material should not only cover word’s meaning, but also all language aspects, such as grammar, subject, and collocation. It is hoped that this paper may offer some useful evidence for English vocabulary (especially synonym) teaching in China.

**References**

*BNC on line*. Retrieved from http://www.natcorp.ox.ac.uk/
*Collins word banks online English corpus*. Retrieved from http://www.ldc.upenn.edu/ldc/online/, http://www.collins.co.uk/corpus/CorpusSearch.aspx
*Longman dictionary of contemporary English*. (2004). Beijing: Foreign Language Teaching and Research Press.
Sinclair, J. (1999). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
Sinclair, J. M. (1987). *Collins COBUILD English language dictionary*. London: Collins.
Sinclair, J. M., Fox, G. et al. (1988). *Collins COBUILD essential English dictionary*. London: Collins.
Appendix A: The Collection of Option and Selection From COBUILD

| Frequent modifiers | Adjectives | Nouns | Verbs |
|--------------------|------------|-------|-------|
| an option           | selection  | option | selection |
| no option           | selection  | option | selection |
| the natural military | wide      | large  | stock  |
| only good           | soft       | final  | call   |
| another jury        | easy       | small  | war    |
| one large           | little     | fine   | latter |
| other his           | attractive | your   | year   |
| soft careful        | viable     | my     | policy |
| best fine           | better     | their  | club   |
| easy team           | preferred  | our    | history|
| this small          | popular    | random | effective|
| first excellent     | nuclear    | careful| co     |
| second extensive    | realistic  | extensive| cost |
| viable sex          | sex        | cheaper| wonderful|
| attractive that      | comprehensive| likely | huge |
| third limited       | cheapest   | fast   | career |
| put pre             | sensible   | represent| zero |
| zero                 | negative   | interesting| risk |
| popular              | huge       | final  | best |

Appendix B: The Collocation of Ill and Sick

| Adverbs               | Nouns          | Prepositions | Frequency (TOTAL 100) |
|-----------------------|----------------|--------------|-----------------------|
| ill                    | sick           | ill          | ill                    | sick          |
| seriously             | mentally       | very         | 3                     | 6            |
| mentally              | severely       | physically   | 2                     | 1            |
| desperately            | terminally     | chronically  | 1                     | 1            |
| terminaly             | severely       | dangerously  | 1                     | 1            |
| severely              | critically     | critically   | 1                     | 1            |
| dangerously           | physically     | physically   | 1                     | 1            |
| critically            | mentally       | mentally     | 1                     | 1            |
| very                  |6                 | during       | ill                    | sick          |
| mentally              |7                 | of           | ill                    | sick          |
| severely              |3                 | to           | ill                    | sick          |
| chronically           |1                 | severely     | ill                    | sick          |
| critically            |1                 | dangerously  | ill                    | sick          |
| mentally              |1                 | critically   | ill                    | sick          |
| animals               | effect                      | at           | about                  | mentally6     |
| bag                   | effects                     | by           | during                 | seriously5    |
| brain                 | feeling                     | for          | of                     | terminally2   |
| building              | health                      | in           | to                     | desperately3  |
| child                 | laughing                    | over         | to                     | severely3     |
| children              | fish                        | with         | ill                    | dangerously1  |
| father                | horses                      | of           | ill                    | critically1   |
| leaves                | lines                       | at           | health                  | 7            |
| lines                 | man                         | during       | treatment               | 3            |
| man                   | pay                         | for          | effect                  | 3            |
| people                | people                      | with         | patients                | 2            |
| wind                  | treatment                   | at           | effect                  | 3            |
| at                    | for                         | 3            | patients                | 2            |
| by                    | with                        | 4            | people                  | 1            |
| by                    | by                          | 1            | people                  | 1            |
| at 3                  | with                       | 3            | people                  | 1            |
| 4                     |                             | 4            | people                  | 1            |
| 5                     |                             | 5            | people                  | 1            |
| 6                     |                             | 6            | people                  | 1            |
| 7                     |                             | 7            | people                  | 1            |