

Research on Corpus-based College English Vocabulary Teaching

Fangmin Pu

School of Foreign Languages, Wuhan Textile University, Wuhan 430000, China.

pfm1@163.com

Abstract. In China, some researchers have investigated corpus-based college English vocabulary teaching and yielded some positive results. Their studies have confirmed the effectiveness of the use of a corpus in EFL vocabulary teaching both in theory and in practice. Nevertheless, as a new approach to vocabulary teaching which can improve the efficiency and effectiveness of EFL learners' vocabulary acquisition, corpus-based vocabulary teaching has not been put into widespread use so far for various reasons such as college English teachers' lack of sufficient knowledge of corpus linguistics and corpus-related software programs. On the basis of reviewing the previous studies on corpus-based college English vocabulary teaching and analyzing the current situation in English vocabulary instruction in China, this paper has offered some useful suggestions on how to carry out further research in this field and how to prompt English teachers to conduct corpus-based vocabulary teaching in EFL classrooms.

Keywords: EFL learners; corpus linguistics; corpus-based vocabulary teaching.

1. Introduction

College English vocabulary teaching is an important part of College English teaching in China. It determines college students' vocabulary acquisition to a large extent, which influences their English development. However, in China, college English vocabulary teaching has not been attached much importance to for several reasons. First, within very limited class time college English teachers need to finish listening, speaking, reading and writing instruction, which consumes most class time, thereby spending only a little time or even no time teaching students' English words and expressions. Second, college English classes are very big and contain students with different English levels, so it's hard to make sure every student really acquires English words the teacher has taught. Third, college students seldom spend time learning English in their spare time, let alone learning English words by themselves. The use of corpora in English classes seems to shed some light on college English vocabulary teaching, for the use of corpora in classrooms can make language teaching and learning highly efficient. In China, some researchers have explored the effectiveness of corpus-based college English vocabulary teaching and have made some achievement in this field. This paper aims to review the studies on corpus-based college English vocabulary teaching in China in recent years and offer some useful suggestions on how further research is carried out in this field and how English teachers put the corpus-based approach to vocabulary teaching into practice.

2. The Definition and Function of Corpus and Corpus Linguistics

A corpus is a collection of machine-readable texts or parts of texts representative of a language or variety, upon which some linguistic analysis can be conducted (McEnery & Wilson, 2001; Meyer, 2004). Corpus linguistics is the empirical research of language which involves the analysis of large databases of texts with the aid of computer technologies (Conrad, 2000). In recent years, methods of corpus linguistics have been employed to explore language-related areas such as language teaching. Many researchers consider corpus-based language teaching much efficient (Johns, 1994; McEnery and Wilson, 2001; Breyer, 2009; Frankenberg-Garcia, 2014; Smirnova, 2017). Many scholars have investigated the effectiveness of the use of corpora in vocabulary teaching. This paper focuses on the research on the use of corpora in college English vocabulary teaching in China.

3. Research on Corpus-based College English Vocabulary Teaching in China

In China, some researchers have investigated corpus-based college English vocabulary teaching and yielded some positive results. Feng & Sun (1999) considered that corpora could provide more reliable and objective language information and help to improve vocabulary teaching. To be more specific, corpora could count word frequency and let EFL teachers know high-frequency words in oral and written English (Feng & Sun, 1999). Feng & Sun (1999) suggested that teachers should help EFL learners acquire those most frequently used words for whether students could use those high-frequency words appropriately would reflect their English proficiency. They further used some examples to demonstrate that if EFL teachers applied corpora to English vocabulary teaching, EFL students could learn word collocation such as the collocation of “over” or “temperature” more efficiently and the differences among synonyms more profoundly (Feng & Sun, 1999).

Liang (2005) introduced the definition of a corpus and its function in language teaching. He also analyzed the theoretical foundation of English vocabulary learning strategies and the relationship between corpora and vocabulary teaching (Liang, 2005). He introduced different types of corpora in the world and stated the possibility of employing them in English vocabulary teaching (Liang, 2005). He further suggested that EFL teachers should construct teaching corpora based on the real teaching needs, and he expounded on how to build a teaching corpus (Liang, 2005).

Huang (2008) explored how to apply the research achievements of corpus linguistics to English vocabulary teaching and suggested that corpus-informed vocabulary instruction (CIVI) should be constructed in terms of teaching syllabus, learning environment and cognitive process. To be more specific, first, feasible vocabulary teaching syllabus should be created on the basis of the analysis of corpus-informed learners’ vocabulary needs; second, Internet-based English learning environment should be created so as to provide learners with sufficient lexical chunks and language samples; third, corpus-informed vocabulary learning cognitive process should be designed to promote vocabulary acquisition.

Zhu (2011) deemed that corpus-based Data-Driven Learning (DDL) was an advanced computer-assisted teaching method, which required EFL learners to observe the authentic English samples provided by corpora and summarize the usage of the target vocabulary which EFL learners need to acquire. This method could stimulate students’ initiative and enthusiasm for English learning, and improve their critical thinking and independent learning (Zhu, 2011). Zhu (2011) proposed the definition and characteristics of DDL, and elaborated on the theoretical foundation of DDL——Constructivism and Lexical Grammar. She also introduced corpus-related software and concordances and stated how to conduct corpus-based data-driven vocabulary teaching in college English classrooms. In other words, there were 4 steps to conduct such a method. First, EFL teachers needed to make sure which word or phrase as well as its usage should be acquired; second, EFL teachers should select a suitable corpus, use a concordances to search the corpus for the search word or node word and classify the concordance lines; third, EFL learners would be asked to observe concordance lines, analyze and summarize the usage of the search word such as its collocation and semantic prosody; fourth, EFL teachers should select and edit some concordance lines into word exercises which EFL learners could do in class to improve their competence in using words appropriately in a certain context (Zhu, 2011).

Long (2014) introduced Sketch Engine and its functions. According to Wikipedia, “Sketch Engine is a corpus manager and text analysis software developed by Lexical Computing Limited since 2003”. Sketch Engine contains 500 ready-to-use corpora in more than 90 languages, each of which has nearly 30 billion words, thereby providing a representative sample of language. Long (2014) stated Sketch Engine mainly has four functions: Concordance, Word Sketch, Thesaurus and Word Sketch Difference. Concordance helps users to get examples of a word form, lemma, phrase, tag or a complex structure; Word Sketch provides users with an automatic summary of a word’s collocations; Thesaurus can show words with similar meanings; Word Sketch Difference compares and contrasts two words with the similar meaning or synonyms by analyzing their collocations. Long (2014) also used specific examples to describe how to use Sketch Engine to improve English vocabulary teaching in detail.

Wang (2012) explored the theoretical foundation of corpus-based English vocabulary teaching and how to use the corpora in vocabulary teaching effectively. He stated that Schema Theory, Lexical Approach and Lexical Grammar could provide the theoretical basis for corpus-based English vocabulary teaching (Wang, 2012). As for Schema Theory, he cited Gui's (2003) remark by saying Schema Theory "is a way in which people use what they have learned to memorize new knowledge." That is, learners could construct language schema to help them acquire new knowledge and enhance their English proficiency, and corpora could provide sufficient English materials for learners to construct and reconstruct that schema (Wang, 2012). Wang also saw Lexical Approach and Lexical Grammar as the theoretical foundation of corpus-based English vocabulary teaching because corpora could help learners comprehend a word's collocation, colligation, semantic prosody and so on, thus contributing to learners' acquisition of lexical chunks both grammatically and semantically (Wang, 2012). Wang reckoned that EFL teachers could use corpora to find out the high-frequency words which are most frequently used in English and concentrate on the instruction of those words (Wang, 2012). Moreover, Wang (2012) expounded on how to conduct corpus-based word collocation, semantic prosody and synonym discrimination teaching.

On the whole, previous studies on corpus-based college English vocabulary teaching have contributed a lot to the field of English vocabulary acquisition. Those studies have confirmed the effectiveness of the use of the corpus in EFL vocabulary teaching both in theory and in practice. However, most studies mainly expounded on the importance and necessity of applying corpora to college English vocabulary teaching and how to conduct corpus-based vocabulary instruction, and very few studies were empirical ones which used experiments or quasi-experiments to test the effectiveness of the use of corpora in college English vocabulary teaching.

4. Current Situation in College English Vocabulary Teaching

Although previous studies have stated and confirmed the effectiveness of the use of the corpus in EFL vocabulary teaching, very few EFL teachers have really conducted corpus-based vocabulary teaching in universities. Currently, most college English teachers adopt relatively conventional approaches to teach EFL students vocabulary in China. For instance, EFL teachers show students some English new words, their English or Chinese meanings and sentence examples containing those words through slides; then teachers ask students to use those words to make English sentences, translate Chinese into English or complete English sentences, etc. If class time is too limited and does not allow teachers to spend much time teaching vocabulary, teachers tend to choose only a few from many new words listed in each unit to explain in detail, and they may even ask students to do vocabulary exercises after class rather than in class.

At present, most college English teachers choose traditional vocabulary teaching method instead of corpus-based vocabulary teaching for the following reasons. First, the majority of college English teachers don't know much about corpus linguistics or corpus-related techniques. As a result, they cannot employ a corpus to assist their vocabulary teaching. Second, even if teachers know how to conduct corpus-based vocabulary teaching, many of them are reluctant to put this new method into practice, because this new method will consume much of their energy and spare time. Third, if teachers intend to conduct corpus-based vocabulary teaching, they need to teach students how to use corpus-related techniques and how to observe the language samples in a certain corpus to find out language rules such as collocation rules. However, it's not easy to let students comprehend those techniques, let alone the skills of observation and induction. Fourth, even if students have been trained to observe and analyze the language samples in a corpus in class, it's very difficult to supervise them and ensure that they really search for the language samples in a corpus and analyze them after class. Therefore, although many scholars have been advocating this new corpus-based approach to vocabulary teaching for a few years, it has not yet come into widespread use.

5. The Application of Corpus Techniques to College Vocabulary Teaching

So far, there have been a number of useful corpus techniques which can be applied to college English vocabulary teaching.

First, if an English teacher intends to conduct corpus-based vocabulary teaching, he or she can construct a particular corpus for that type of teaching. Prior to the construction of the corpus, he or she needs to consider some crucial factors related to the creation of the corpus: the length of the corpus, the genres of texts which will be included in the corpus, the length of the individual texts, the number of the text samples, the educational level, gender and age of speakers and writers included in the corpus and the like. Once how to construct the corpus has been planned, the teacher needs to embark on the actual creation of the corpus which involves three time-consuming steps: collecting, computerizing and annotating the texts included in the corpus. As regards annotating a corpus, the teacher needs to learn how to use tagging and parsing tools available to tag and parse the corpus. If teachers are unwilling to construct corpora by themselves, they can use the corpora available online such as BNC and ANC.

Second, teachers need to learn some useful software tools to conduct corpus analysis and corpus-based teaching. There are a number of software programs that can retrieve much useful information from corpora: The Linguistic Data Base (LDB) for analyzing the Nijmegen Corpus; ICECUP for analyzing the British component of ICE; Sara to analyze the British National Corpus; and different concordancing programs (Meyer, 2004). Concordancing programs such as WordSmith and AntConc have a variety of functions: “to conduct simple searches for words, groups of words, or words containing prefixes and suffixes; to display the results of a given search in KWIC (key word in context) format; and to calculate the frequency of the item being searched in the corpus in which it occurs” (Meyer, 2004). After teachers acquire how to employ those software programs, they can display the data extracted from the corpus used to facilitate vocabulary teaching in class and ask students to observe and analyze the corpus data which contain the target vocabulary, or give out handouts which contain the language samples extracted from the corpus to students and ask them to observe and analyze the samples after class. Admittedly, teachers also need to train students how to analyze the samples from a corpus and take measures to ensure students really make observations and analysis of the handouts after class to make the corpus-based vocabulary teaching much effective.

6. Conclusion

In China, corpus-based EFL vocabulary teaching is a new methodology which may promote EFL learners' vocabulary acquisition. Some scholars have conducted empirical and theoretical research to confirm the effectiveness of corpus-based vocabulary teaching, and thus have advocated that the corpus-based teaching method needs to be introduced into English teaching such as English vocabulary teaching in universities. However, very few teachers have really applied a corpus to their vocabulary teaching for a few reasons such as their ignorance of corpus-related techniques. Therefore, if we want corpora to be put into widespread use, we need to encourage English teachers to acquire corpus-related software programs first, which include concordancing programs such as WordSmith and AntConc. In addition, English teachers need to realize the significance of teaching vocabulary and really attach importance to it, thereby trying different ways to enhance vocabulary teaching such as conducting corpus-based vocabulary teaching. In the future, researchers need to conduct more experiments to prove the effectiveness of using corpora in vocabulary learning and teaching so that English teachers will have more confidence in putting the new approach to vocabulary acquisition into practice.

References

- [1]. Breyer, Y. Learning and teaching with corpora: reflections by student teachers. *Computer Assisted Language Learning*. Vol.22 (2009) No.2, p. 153 – 172.

- [2]. Conrad, S. Will corpus linguistic revolutionize grammar teaching in the 21st century? *TESOL Quarterly: TESOL in the 21st Century*. Vol.34 (2000) No. 3, p. 548 – 560.
- [3]. Feng, Yuejin & Sun, Mingyu. On the role of a corpus in English vocabulary teaching. *Foreign Language Research*. (1999) No.3, p. 60-62.
- [4]. Frankenberg-Garcia, A. The use of corpus examples for language comprehension and production. *ReCALL*. Vol.26 (2014) No.2, p. 128 – 146.
- [5]. Gui, Shichun. Memory and English learning. *Foreign Language World*. (2003) No.3, p. 2-8.
- [6]. Huang, Ruoyu. The construction of corpus-informed vocabulary instruction model. *CAFLE*. (2008) No.2, p. 52-57.
- [7]. Johns, T. From printout to handout: grammar and vocabulary teaching in the context of data driven learning. In: Johns T, *Perspectives on Pedagogical Grammar*. New York: Cambridge University Press, 1994.
- [8]. Liang, Sanyun. Research on corpora and vocabulary teaching strategies. *CAFLE*. (2005) No.5, p. 33-36.
- [9]. Long, Yu. The application of corpora in English vocabulary teaching. *English Square*. (2014) No.6, p. 67-70.
- [10]. McEnery, T., & Wilson, A. *Corpus linguistics*. Edinburgh: Edinburgh University Press, 2001.
- [11]. Meyer, C. F. *English corpus linguistics: An introduction*. Cambridge: Cambridge University Press, 2004.
- [12]. Smirnova, E. A. Using corpora in EFL classrooms: the case study of IELTS preparation. *RELC Journal*. Vol.48. (2017) No.3, p. 302 – 310.
- [13]. Wang, Jiayi. Corpus-based English vocabulary teaching: theory and application. *Foreign Language Research*. (2012) No.4, p.127-130.
- [14]. Zhu, Huimin. Data-driven learning: the new trend of English vocabulary teaching. *CAFLE*. (2011) No.1, p. 46-50.