



BNC-Based Design of College English Vocabulary Teaching for Chinese College Students

HUO Yanjuan^{[a],*}

^[a]School of Foreign Languages of China West Normal University, Nanchong, China

*Corresponding author.

Received 12 February 2014; accepted 26 May 2014

Published online 26 June 2014

Abstract

The present study aims to bring about some innovations in construction and designing of vocabulary teaching on the basis of the BNC (British National Corpus). Corpus linguistics, providing the enormous and authentic language material, has been widely employed by scholars, teachers and researchers in various fields and perspectives of language research. Therefore, it has made significant contribution to language theory construction and gives a lot of inspiration to English foreign language (EFL) teaching. The present research applying the BNC corpus to the vocabulary teaching is intended to effectively improve teacher's input and introduce corpus-based approach. Moreover, it wants to show that the computer-aided curriculum designing can effectively facilitate students' vocabulary learning and improve their language proficiency.

Key words: BNC; Corpus; EFL

Huo, Y. J. (2014). BNC-Based Design of College English Vocabulary Teaching for Chinese College Students. *Studies in Literature and Language*, 8(3), 122-125. Available from: <http://www.cscanada.net/index.php/sll/article/view/4960>
 DOI: <http://dx.doi.org/10.3968/4960>

INTRODUCTION

As Wilkins claimed that without grammar little can be conveyed and without vocabulary nothing can be conveyed. In the process of second language acquisition, vocabulary is fundamental and critical which will influence the quality and quantity of the second language

acquisition directly and indirectly, yet most Chinese students often complain that vocabulary acquisition is a tedious and tough task. Corpus, as an assistant tool in language education, has been proved to have great potentials in facilitating language teaching and learning. Comparing with the achievement of applying corpus to real classroom teaching over the world, the application of corpus in English foreign language teaching within Chinese context seems less fruit. The present study aims to promote the wide use of corpus in designing English teaching in China. It will start from the very basic knowledge of BNC corpus. Detailed examples including differentiation of English synonyms, lexical chunks and development of the communicative competence of students will be given with the purpose for making our teaching effective and efficient. Demonstrations of how the corpus based technology can be used in language teaching will also be included.

1. BASIC KNOWLEDGE OF BNC CORPUS

BNC (British National Corpus) which is made up of over 100 million word collection of samples contains both spoken and written texts. Kennedy pointed out that the BNC, which was built within four years, was designed to be well balanced and to be widely accessible for educational, academic and commercial purposes. The BNC includes 90% written word collection of samples and 10% spoken part. The written part has a wide range of sources including source extracted from regional and national newspapers, specialist periodicals and journals, popular fiction and academic books etc.. The spoken language is collected from different contexts consisting of orthographic transcriptions of unscripted informal conversations. All these samples of written or spoken language are for different ages and various interests. To make the introduction of BNC concise and vivid, the following table is used to show the basic indexes of BNC.

Table 1
BNC

Feature	Size (million of words)	Scale of spoken language (million)	The span of time	Update frequency per year (million)	Dialect	Availability
Figure	100	10	1970-1993	0	British	Free / web

2. APPLICATION OF BNC CORPUS IN COLLEGE ENGLISH VOCABULARY TEACHING

Teachers could provide a brand new approach to teach English vocabulary by combining BNC corpus with traditional teaching style, which learners are enabled to grasp the subtle difference of the synonyms in different context and to gain awareness of the characteristics of language variety with a great number of authentic English material provided by BNC corpus.

2.1 Differentiation of English Synonyms

In terms of English vocabulary acquisition, students are not only supposed to learn the largest amount of words, but also make clear the difference of numerous near synonyms which make Chinese students feel tougher to learn and remain a big obstacle to be overcome. According to Lyons, synonyms are expressions with the same meaning, whereas near-synonyms are expressions that are more or less similar, but not identical in meaning. It is hard for Chinese students to differentiate English synonyms which occupy a great number in English language and clear differentiation and proper use of synonyms are the critical requirement for students to master the accurate use of vocabulary and keys to improve the language learning skills. Traditionally teachers try to identify the discrimination of synonyms by relying on the intuition and introspection but as we know, it is not a scientific and objective teaching method. Since the booming of computer technology and the emergence of corpus, corpus-based approach as a brand new teaching approach in applied linguistics is more and more popular and welcomed by English teachers. Based on BNC corpus, the author differentiates the sampled synonyms from genre/frequency distribution. By analyzing the data and tables of synonyms the author is intended to illustrate a more persuasive way to distinguish synonyms in English thus help Chinese students get an effective and clear distinction of synonyms. If a teacher wants to show the difference of the near synonym “accurate” and “right” to students, they can firstly check the chart list of these two words. The following table is the checking result:

The Table 2 shows us that as synonyms, “accurate” and its synonyms “right” have a lot of difference in using style and genre distributions and their frequency is also

different. By observing the frequency/genre distributions of “accurate” and its synonyms “right” in BNC, teachers can show the basic difference and help students try to identify the appropriate use of these two words. In order to provide more believable and persuasive teaching and facilitate student’s understanding of the great number of synonyms, it is important for teachers to integrate traditional teaching method with corpora-based analysis. If some students are given the above table they will feel it is directive and persuasive to distinguish the synonyms.

Table 2
The Result of BNC Chart Checking

Genre	Accurate		Correct	
	Frequency	Per million	Frequency	Per million
Spoken	161	16.16	740	74.27
Fiction	187	11.75	533	33.50
Magazines	215	29.61	584	80.42
Newspaper	176	16.82	354	33.82
Academic	731	47.68	1693	110.43
Total	2851	29.62	6584	68.40

2.2 Lexical Chunks Instruction Based on BNC

Chinese EFL learners study English in the classroom settings so that the input of teachers has significant impact on their target language proficiency. The construction of teaching curriculum should play an important role in helping students to acquire the native-like fluency. In 1956, psychologist Miller first proposed the concept of “chunk” and he gave the definition of chunk as the smallest unit of people’s short-term memory in which people were most familiar. It is important for teachers to take familiarity into consideration because familiarity plays a crucial role in fluency of target language output. If lexical chunks have already stored, second language learners will feel familiar with language form and it can save a lot of time and effort and speed up language processing. Students can use lexical chunks as the framework or structure of the sentence. Let us take the lexical chunk “it is + adj” for example. The first sixteen result of data extraction from BNC corpora is shown in the following Table 3.

Table 3
The Search Result (Partial Listing)

Context	Frequency	Number	Context	Frequency
It is possible	3,353	9	It is impossible	807
It is important	3,008	10	It is easy	770
It is difficult	1,906	11	It is hard	761
It is clear	1,769	12	It is unlikely	651
It is necessary	1,399	13	It is interesting	610
It is likely	1,141	14	It is better	381
It is true	1,104	15	It is vital	365
It is essential	828	16	It is good	331

Table 3 shows frequency information of the first sixteen lexical chunks which have the structure of it

is + adj. Through the observation of the distribution of this structure, we can find the most frequent lexical chunks belonging to this structure used by native speakers. BNC corpus can provide learners with a huge number of this pattern. Once clicking the button of search we can get 100 expressions in context part and students and teachers can get more by clicking the button of “see more than 100 words and phrases”. With rich resources of naturally occurring authentic language, BNC can help teachers to check lexical chunks against intuition. The Table 3 only lists the first sixteen lexical chunk of “it is adj” and it is easy to discover more lexical chunks if learners want to know. If we check more, we can get: it is amazing; it is misleading; it is unrealistic; it is unwise; it is inconceivable; it is extraordinary; it is useful; it is probable; it is significant; it is reasonable; it is advisable; it is available; it is surprising; it is appropriate, etc. With the great number of lexical chunks offered by BNC corpus, students will expose themselves to authentic input before they are able to speak so that students have more potential in applying various lexical chunks to their English output. When these lexical chunks are consolidated through memorization, they will be reserved in learners’ memory and become available for students’ future use. Teachers should also give suggestions to students that the more practice and memorization the better because the greater efforts in storage and more active practice will make vocabulary solid in long-term memory and students will get familiar with their flexible use. Lexical chunks should gain their deserved attention by teachers because they make their great contribution to improving students’ language proficiency.

2.3 Develop the Communicative Competence of Students

2.3.1 Cross-Culture Communication

Teachers in China are encouraged to cultivate students’ communicative competence since the issue of Chinese College English Curriculum Requirements (Requirements) in the year of 2004. Because of the Requirements, many English teachers are asked to develop students’ ability to use English in an all-round way, to enhance the language proficiency of students and to cultivate the cultural sense of students so as to meet the needs of social development and international exchanges of China. Therefore, the whole acquisition of a target language does not only mean the apprehension of grammatical rules but also the communicative competence. Chinese college students have encountered great problems in correctly using words in real communication, their limited lexical knowledge and little sense of cultural difference constituting an obstacle to achieve the native like communicative competence. That is why a long-term goal of language teaching gives the importance of cultivating cross-culture

communication ability. It becomes an unspoken agreement that language and culture are both very important to facilitate the language proficiency of students. However, deep-rooted by mother tongue, Chinese students are easily using chinglish or word to word translation which makes foreigners feel confused and are seldom used by native speakers. For example, Chinese students tend to say “learn knowledge” which is obviously influenced by Chinese expression of “study” and “knowledge” whereas native speakers use “acquire knowledge” to express the same meaning. Facing these problems, students should be told that the acquisition of a word means that students are quite familiar with words habitually used with other typical words in the typical context rather than been separated or isolated. Aiming at improving cultural sense of students, teacher can design some vocabulary related activities and practice, showing target vocabulary with its phonetic, semantic as well as pragmatic knowledge repeatedly and consciously. In order to educate students’ communicative competence, teachers are also supposed to offer an abundance of real language material in various context and situations and the huge number of empirical information which gives great challenging to teachers. Thanks to the rapid development of China and the decrease of prices of computers, most universities in China are equipped with the advanced technology of computers and have enormous chances to the Internet so teachers can get accessed to the computer-assisted and corpus-based instruction of English with elaborating the design of the learning environment based on the combination of corpus and intuition and teaching experience of teachers. Corpus, as we know, can provide teachers and students rich contexts that are collected from texts or speeches by means of random sampling. The potential of the application of corpora in teaching as a target language is tremendous not only because it is powerful tools to the change of the traditional teaching model but also because the application of data-driven learning which is based on corpus can provide teachers a great number of native expressions. Thus, the application of corpora in English teaching is of great significance of both the construction of teaching resources and the reform of teaching.

2.3.2 KWIC Searching

The word corpus, derived from the Latin word means body which can provide precious authentic language. KWIC is a short term for key word in context. With the help of the tool of KWIC, we can investigate language structure including sentences, paragraphs and discourse by highlighting the keywords. Teachers can analyze large databases of real language examples stored on computer. By checking “The Great Wall” as keyword in context display, we can get a word list generated from BNC which is characterized by a text or a genre. Teachers can obtain the expanded context by clicking the keyword list. These following are two expanded contexts generated from BNC.

2.3.2.1 Expanded Context 1

Mapping the universe. Astrophysics Piyush Ojha. MARGARET Geller and John Huchra of the Harvard-Smithsonian Centre for Astrophysics have discovered a thin sheet of galaxies measuring at least 200 million by 500 million light years, which stretches across the sky like a giant canopy. They call this structure 'The Great Wall'. Its existence casts doubt on the current astronomical theory of the distribution of galaxies and the large-scale structure of the universe. The observation of the Great Wall has come from an ambitious project to map the positions of all visible galaxies which are brighter than a specified minimum. The survey –'mapping the universe' -- which Geller and Huchra began a few years ago is not yet complete but they have seen enough to conclude that 'the size of the largest structures we detect is limited only by.

2.3.2.2 Expanded Context 2

For, as the speed of light rushed down upon her, she heard, in the pause between the upbeat of Fenna's wings and their down swoop, she heard the singing of the spheres, the harmony of the universe which turns, turns, turns in the atom and in the infinite and is unmoved in its own movement and is all one single force. Now, in the dream influence of the moon, Fenna told Maggie to look down on all the kingdoms of the world, and across China she saw the snaking of the Great Wall. On the southern side she saw the mandarin silks of the beginnings of civilisation and the huge courage of the Long March; and on the other, to the north, the eternal silence of the greatest desolation. She looked on north, north, north, through desert and tundra, to the cruel unbroken peace of the polar seas. And Fenna awaited her thanks for a journey so beautiful and so unique, the gifts that only he could offer her.

The first expanded context is quoted from Guardian Newspapers Limited in the year of 1989 in London while the second expanded context is selected from the fiction and prose part provided by Chatto & Windus Limited in the year of 1990 in London. By analyzing and observation of keywords in context, students can find the characteristics of the keyword in context written by native speakers. Students can reveal the salient feature of Chinese cultural symbols such as The Great Wall, Mandarin silks, the Long March used by some speakers comparing with a metaphor used by Astrophysics Piyush Ojha. MARGARET Geller and John Huchra. They compare the structure of a thin sheet of galaxies with 'The Great Wall'. Therefore word has a close relationship with the context they occurring. The major findings of these two extended contexts can provide certain pedagogical implications, such as how particular a word behaves in different contexts and how different or similar the same word in different registers. With banks of electric texts provided by corpus, students themselves will become aware of collocation, semantic and

pragmatic English expressions gradually and naturally. Meanwhile, the great number and range of native language input will consequently make Chinese students recall unconsciously the typical context and better distinguish their potential environment. From words to sentences then to discourse, with the aid of a corpora and KWIC tool, students will be displayed a progressive acquisition of the words immersed with native language and culture instead of teaching separate words without real life contexts. The corpus-based teaching approach can help Chinese learners choose flexibly the appropriate word in the real communication and facilitate their communicative competence.

CONCLUSION

This paper constitutes an attempt to explore the corpus-based teaching approach of lexical competence development on the part of non-English majors in China. Through the corpus work and data analysis, this paper has made exploration into the argument of how to improve English vocabulary teaching and bring about innovative and academic values. Near synonyms with character of overlapping in some core parts of their meaning but difference in detail can be distinguished by corpus-based method. For lexical chunks, the above study intends to improve teacher's input by combining traditional textbook-based teaching style with utilizing computer-aided method. The Internet resource can provide Chinese students numerous lexical chunks thus facilitate student's native like output. Moreover, the large BNC corpus with abundant information from vast, authentic language 'body' can provide a scientific basis for teachers to choose and enriches their teaching content. It is an ideal way to meet the requirement of developing communicative competence of students and cultivating their language proficiency. All in all, a combination of traditional teaching style and corpus-based teaching is suggested to be applied in vocabulary teaching and in this way some improvement could be done in the future.

REFERENCES

- BNC (The British National Corpus). (1970-1993). Retrieved from Website <http://corpus.byu.edu/bnc/>
- Kennedy, G. (2000). *An introduction to corpus linguistics* (p.50). Beijing: Tsinghua University Press.
- Lyons, J. (1995). *Linguistic semantic: An introduction* (p.60). Cambridge: Cambridge University Press.
- Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, (63), 81-97.
- Wilkins, D. A. (1972). *Vocabulary Linguistics in Language Teaching* (p.109). London: Edward Arnold Ltd.