

# International Journal of Social Science and Education Research



ISSN Print: 2664-9845  
ISSN Online: 2664-9853  
Impact Factor: RJIF 8.00  
IJSSER 2024; 6(1): 90-94  
[www.socialsciencejournals.net](http://www.socialsciencejournals.net)  
Received: 15-12-2023  
Accepted: 20-01-2024

**Zhang Miaomiao**  
Graduate Student, College of  
Foreign Languages, Yangtze  
University Jingzhou, Hubei  
Province, China

**Corresponding Author:**  
**Zhang Miaomiao**  
Graduate student, College of  
Foreign Languages, Yangtze  
University Jingzhou, Hubei  
Province, China

## A study on vocabulary teaching in senior high schools with the aid of corpora

**Zhang Miaomiao**

**DOI:** <https://doi.org/10.33545/26649845.2024.v6.i1b.84>

### **Abstract**

Corpus has been increasingly emphasized by the educational community for its three major advantages of being able to accurately count word frequencies, provide authentic and rich language materials, and create authentic and diverse language situations. With the continuous creation, development and updating of corpora and corpus retrieval tools, the research methodology and technological means of corpora will be continuously applied to the practice of high school English vocabulary teaching, and this paper analyzes the different applications of corpora to assist high school English vocabulary teaching and the teaching implications. This paper analyzes the different applications of corpus to assist high school English vocabulary teaching and the teaching revelations.

**Keywords:** Corpus, teaching English vocabulary, word frequency, language materials, language contexts

### **1. Introduction**

Vocabulary is one of the 3 major elements of language, the pillar on which the language system depends. However, for a long time, our junior high school and senior high school English teaching has been organized and implemented with grammar, sentence pattern or function as the main focus, and English vocabulary teaching has often been in a subordinate position to teaching. Even in the process of teaching vocabulary alone, it is easy to fall into the misunderstanding of holding the word list and memorizing it by rote. It is gratifying that in recent years, with the deepening of our English teaching reform, English vocabulary teaching has received more and more attention from the domestic educational circles, who have carried out extensive and in-depth research and discussion on it, and achieved fruitful research results, such as the use of cognitive psychology, cognitive linguistics, semantics, stylistics, pragmatics, cross-cultural communication, and corpus, etc., to carry out research on the teaching of English vocabulary from a variety of perspectives. Research. Among them, due to the rapid development of computer science and network technology, rich network resources also make foreign language teaching faster and more convenient, computer linguistics and corpus linguistics have also developed rapidly, and corpora provide reliable and objective corpus for vocabulary teaching with their representative macro database, and the introduction of network corpora in teaching has become a new trend in the development of foreign language education today.

A corpus is a linguistic database consisting of a large amount of information about the actual use of a language, dedicated to the study, analysis and description of language. A corpus is created by collecting people's actual use and representative real language materials on the basis of random sampling modern corpus linguistics (Modern corpus linguistics) is a science of language research in the middle and late 20<sup>th</sup> century. It is based on a large amount of real language materials, and starts from investigating the distribution frequency of language information to study the laws and patterns of language in real use. At present, there are different corpora in different countries of the world, and the linguistic information and materials they provide for linguistic research are not the fabricated words of some individuals' imagination, nor are they simply a few or even individual examples extracted from some books, but a large number of real-life spoken and written words used by the speakers of English as their native language are collected in a systematic way with a purpose, and this builds up the description of the language on the basis of a vast and authentic corpus

(The real corpus is the base of the language study) authentic and reliable corpus. This not only improves the accuracy and reliability of our language research, but also helps us in English teaching. In English teaching, vocabulary teaching is one of the earliest and most fruitful areas in which corpus resources and research tools have been applied to foreign language teaching (Tang Jieyi, He Anping, 2004) <sup>[33]</sup>.

## 2. The application of corpus in teaching English vocabulary in high school

Corpora provide extremely rich teaching resources and brand new research ideas for vocabulary teaching. The corpus-based vocabulary teaching method is in line with the student-centered teaching concept: teachers are no longer the transmitters of knowledge, but more importantly the organizers, designers and guides of teaching activities, and students are the researchers and participants, explorers and discoverers. Utilizing corpus can not only help students to better master knowledge, but also improve their abilities in many ways. That is to say, it helps us to change our concepts, absorb and accept teaching concepts and methods that are different from those of the traditional teaching mode, and promote a new change in the mode of teaching English vocabulary, which can be broadly categorized into the following aspects.

### 2.1 Statistical analysis of word frequency using the corpus

Vocabulary teaching is important in high school English teaching, whether it is for test-taking or developing students' quality. But in the real teaching process, it is difficult for teachers to utilize the limited time to improve students' vocabulary effectively and efficiently. Some teachers present all the words to the students and give them a list of words to memorize without explanation or focus, which is half the result with twice the effort. Some teachers lack an overall understanding of the vocabulary used in the textbook and do not have a full grasp of the vocabulary layout and difficulty of vocabulary in the textbook, which makes it difficult to organize vocabulary teaching effectively. In addition, the vocabulary list often lists some low-frequency rare words without selecting them, while some common and high-frequency words outside the vocabulary list are often ignored by teachers, so that students' vocabulary is insufficient, vocabulary use is of low quality, and vocabulary structure is unreasonable, so there is a need to innovate vocabulary teaching.

The statistical analysis of word frequency based on corpus and then targeted teaching can be well adapted to this demand and can be a good solution to this problem.

Utilizing corpus for statistical analysis of word frequency can determine the range and number of high-frequency vocabulary of different levels more accurately, and in foreign language teaching, corpus word lists can provide an objective basis for determining the range of vocabulary in the syllabus as well as the selection criteria. Teachers can utilize common corpora such as Range and COCA to achieve this purpose when teaching. In addition, teachers can also use the corpus to analyze the frequency of words and then prepare word lists that are more suitable for students' learning. That is to say, in the actual teaching process, teachers can use the existing corpus to prepare targeted personalized word lists according to different teaching objects, different teaching purposes and different

teaching stages, which is not only suitable for one-to-one teaching, but also for class teaching and hierarchical teaching. In China, JDEST corpus has been one of the main bases for the formulation and revision of word lists of university English syllabus for many years, and the Naval College of Aviation Engineering has relied on the information sources such as JDEST Academic English Corpus and English National Corpus, and screened and identified 3,000 repetitive vocabularies as the key points of vocabulary teaching in conjunction with its own teaching purposes, which has received remarkable teaching effects (Wang Jing, 2005) <sup>[35]</sup>.

In vocabulary teaching in high school English, teachers should first assess the difficulty and distribution of vocabulary in the textbook, determine the key vocabulary, and then refer to the teaching based on the real corpus of the Corpus of Native English Speakers. The current English syllabus for high school requires a vocabulary of 3,500 words, but in real teaching, it is unrealistic to divide the time equally among these 3,500 words, and teachers should devote their limited time and energy to presenting the most commonly used words. If high school students can first master these two thousand or so high-frequency vocabulary words, they will be able to get a basic understanding of the general idea of a text. Through the corpus, we can then learn very well what words are high-frequency and what words do not need to be memorized in so much detail. For English vocabulary teaching, high-frequency word lists are extremely valuable, on the one hand, they can help us to determine the content of vocabulary teaching, find out the focus of teaching, and even arrange the order of teaching, and provide various effective references for both teachers and learners; on the other hand, in combination with high-frequency word learning, the frequency of exposure to key vocabulary of the students can be increased, which can also improve the students' self-confidence and learning efficiency. Xiang Zhenzhen's (2015) <sup>[43]</sup> use of Range for college English vocabulary grading teaching provides a good example for us. In high school English teaching, we can similarly classify the word lists listed in the textbook into difficulty levels corresponding to Range15 level word lists, and refer to the High School English Curriculum Standards (Revised in 2020) to classify the 15 level word lists into five levels of difficulty, and then teach according to the arrangement of difficulty and word frequency, starting with the key vocabulary and target vocabulary, and appropriately reducing the bias towards non-key vocabulary. In other words, teachers can use the corpus for word frequency analysis and statistics to come up with a high-frequency vocabulary bank, and these really commonly used words are the main content and key targets that can be identified in the process of reading and learning, and our teaching focus should be placed on the linguistic patterns related to the high-frequency words, so as to alleviate the vocabulary burden of the students in reading.

### 2.2 Teaching word collocation using a corpus

Vocabulary collocation (Collocation) is one of the most important concepts and research contents in the field of linguistics. To learn a word, not only to learn its meaning and pronunciation, but also different collocations are quite important, which is a key and difficult point for vocabulary teaching of English teachers in China. And by using corpus

to teach word collocation, students can learn the typical collocation of words intuitively.

Vocabulary learning activities based on contextual co-occurrence in the corpus are designed so that teachers can learn the high-frequency collocations of the target vocabulary by utilizing online corpora such as BNC, where teachers can have quick and free access to these real-life examples of collocations, which can help students to build up a notion of the frequency of collocations by understanding the tendency of native speakers to use certain collocations. In traditional teaching, teachers tend to present tools such as dictionaries to teach students some target collocations directly, while the frequency of using English vocabulary revealed by the corpus reflects the tendency of native speakers to use vocabulary, and it is also the authentic and standardized expression collocations, which is very objective and reliable. Therefore, using the corpus to learn vocabulary collocations can not only verify whether our collocations are the same as those of native speakers, but also avoid the negative transfer effects of the mother tongue.

Through the corpus-based word collocation research, we can break through the limitations of phrases and obtain more comprehensive information about word collocations on the one hand; on the other hand, we can remove those outdated and useless collocations and learn real-life high-frequency co-occurring words (collocations), reduce the arbitrariness and limitations of presenting collocations in vocabulary teaching and accelerate the comprehensive mastery of collocations and revitalization of vocabulary learning. Activation, improve the quality and efficiency of vocabulary learning; in guiding students to observe contextual co-occurrence and learn vocabulary collocations, it can also improve students' learning ability and thinking quality.

### 2.3 Teaching thematic vocabulary using a corpus

The English Curriculum Standard for General High Schools (Revised 2020) points out that vocabulary learning at the high school level is not mere memorization of words, but more importantly, it is to perceive and understand the meaning of relevant topics in discourse, and to use words to express information and meanings of the relevant topics through listening, speaking, reading, reading, and writing (Ministry of Education, 2021). It can be seen that the perception of vocabulary in high school level is built on the basis of thematic contexts. It also suggests that teachers should guide students to understand the meaning of words by using the structure of the words and the context of the text, and to learn the usage of the words with the help of resources such as dictionaries; in other words, words should be taught in the context of the discourse, and in the context of comprehension, rather than solitary word memorization. In addition to the use of frequency for hierarchical teaching, vocabulary matching, etc. mentioned above, another characteristic use of the combination of corpus and vocabulary teaching is the use of corpus to extract key words so as to further understand the discourse.

Teachers can use AntConc or WordSmith tools to compare the text to be analyzed with the common corpus and extract the key words automatically. Therefore, when students encounter unfamiliar vocabulary, teachers can guide students to observe the contextual co-occurrences to guess the words, and such guessing game activities not only connect the topics, but also enhance students' motivation

and learning fun. In addition, the teacher can find more directions of topics through these keywords, which can further introduce the topic and improve students' participation and motivation in the classroom. In addition, teachers can also use the plot function to present the distribution of the keywords WE, and guide students to analyze the type and content structure of the discourse according to the position of the keywords in the text, in which the students' thinking about the article is not only limited to the thematic content of the article, but further shifted to the genre type of the article.

To sum up, we can summarize the steps of using corpus in teaching thematic vocabulary: first step, the teacher extracts the key words of the text by using the corpus software, guides the students to guess the content of the article, and lets the students develop their imaginations to expand the content of the theme words, imagining the topics that the article is likely to expound on and the contents they contain. In the second step, the teacher clicks on the keywords, presents the contextual co-occurrence of the keywords, and guides students to further explore the specific content of the text, which is different from traditional teaching in that the teacher always acts as a supporting role, and her role is to help students discover the problem step by step by themselves, and to get the pleasure and sense of accomplishment of exploring from it. The key to the process of guiding is not to analyze the words and sentences in the common current step by step, but to let the students through the words, linking sentences, and ultimately linking the discourse, to achieve the effect of "understanding the words in the discourse, through the words to understand the discourse". In the third step, the teacher helps the students to clarify the thoughts of the article through the key words, construct the cognitive schema of the discourse structure and familiarize themselves with the style of the discourse.

Unlike the traditional English classroom, teachers should introduce corpus-assisted teaching tools into the classroom, teach the textbook and the corpus effectively together, and verify and consolidate what is taught through corpus tools. Led by the thematic context, students' extended and deepened learning of thematic vocabulary is strengthened by designing a large number of language learning activities, which ultimately leads to the internalization of thematic vocabulary. In other words, teachers can use the corpus to guide students to expand their vocabulary and deepen their learning under the leadership of the thematic context, further strengthen their understanding and use of vocabulary, and realize the two-way facilitation of comprehension of discourse and words.

### 3. Discussions and Conclusion

To summarize, the introduction of corpus in high school English vocabulary teaching can improve the teaching effect from multiple dimensions. Flexible use of corpus to assist vocabulary teaching, its ways and functions are not only limited to the several aspects described by the author in the article, so in the future research can change the idea, continue to explore, explore. This is not only a change in the form and means of teaching, but also a revolution in the thinking and methods of foreign language teaching. It has profoundly changed China's long-standing teacher-centered traditional teaching mode, and further established a student-oriented, proactive and personalized teaching method. In conclusion, corpus with its multidisciplinary, multifaceted

and multifunctional advantages is increasingly being emphasized by the educational community. With the continuous creation, development and updating of corpus and corpus retrieval tools, the research methodology and technological means of corpus will be continuously applied to the practice of English vocabulary teaching, and the research on English vocabulary teaching based on corpus will also be continuously developed and deepened.

#### 4. References

- Cheng W, Warren M, Xun-Feng X. The language learner as language researcher: Putting corpus linguistics on the timetable. *System*. 2003;31(2):173-186.
- Conrad SM. The importance of corpus-based research for language teachers. *System*. 1999;27(1):1-18.
- Conrad S. Register in corpus linguistics: the role and legacy of Douglas Biber. *Corpus Linguistics and Linguistic Theory*, 2022, (1).
- Flowerdew L. Corpus linguistic techniques applied to textlinguistics. *System*. 1998;26(4):541-552.
- Xia G. The Application and Research of Corpus Method in College English Teaching Reform. In: *Proceedings of 2019 International Conference on Humanities, Management Engineering and Education Technology (HMEET 2019)*. Clausius Scientific Press, Canada; c2019. p. 372-381.
- Henry A. *Small corpus studies and ELT: theory and practice*. Sage Publications; c2001.
- Natalia G. *International Corpus of Learner English: Implications for ELT*; c1998.
- Johns T. Data-driven learning: the perpetual challenge. In: Kettemann B, Marko G, eds. *Teaching and Learning by Doing Corpus Analysis*. Amsterdam: Rodopi; c2002. p. 107-117.
- Lewis M. *The Lexical Approach: The state of ELT and the way forward*. Hove, England: Language Teaching Publications; c1993.
- McCarthy M. *Issues in Applied Linguistics*. Beijing: World Affairs Press; c2006.
- Sinclair J. *Reading Concordances*. London: Pearson Education Limited; c2003.
- Wang D. The Study of Corpus as a Tool to Promote the Development of English Teaching. In: *Proceedings of the 2014 Conference on Informatization in Education, Management and Business*. Atlantis Press; c2014. Doi:10.1515/CLLT-2022-0032.
- Chen JS. Corpus Linguistics and English Language Teaching. *Journal of PLA College of Foreign Languages*. 2004(01):60-64+127.
- Chen Y. Corpus linguistics and English vocabulary teaching. *Higher Education Forum*. 2006;(02):66-70.
- Feng YJ, Sun MY. On the role of corpus in English vocabulary teaching. *Foreign Language Research*. 1999;(03):60-62.
- Gui SC, Yang HC. *Corpus of Chinese English Learners*. Shanghai: Shanghai Foreign Language Education Press; c2003.
- Han M, Xin X. Corpus and Vocabulary Teaching in English Reading Classes. *Journal of Xi'an International Studies University*. 2009;(04):105-107+112. doi:10.16362/j.cnki.cn61-1457/h.2009.04.024.
- Ho AP. Talking about corpus research. *Foreign Languages (Journal of Shanghai International Studies University)*. 1997, (05).
- He A. A study of vocabulary knowledge development in English language textbook series: A dynamic systematics perspective. *Foreign Language Teaching and Research*. 2015(06):898-908+961.
- Huang P, Yan C. Analysis of the current situation of domestic corpus application in English teaching research. *Shandong Foreign Language Teaching*. 2010(05):44-50. doi:10.16482/j.sdwy37-1026.2010.05.020.
- Kong L, Qin H. The use of corpus in vocabulary teaching: Lexical stratification and instructional design. *Foreign Language Teaching Theory and Practice*. 2013(04):58-63+95.
- Kong W. Corpus-based design of high school English vocabulary teaching. Master's thesis, Yangzhou University; c2012.
- Lee CW. Corpus-driven English vocabulary teaching: Rationale and application - A review of How to utilize corpus for language teaching. *Chinese College Technology (Z1)*; c2023. p. 146. doi:10.16209/j.cnki.cust.2023.z1.023.
- Liang S. A study of corpus and vocabulary teaching strategies. *Foreign Language e-Learning*. 2005, (05).
- Linsen. Cultivating Students' Core Literacy in English Discipline in Vocabulary Teaching Based on Corpus. *English Teacher*. 2020;(23):187-192.
- Liu M. The application of corpus in English vocabulary teaching: Patterns, gains and losses, and prospects. *Crazy English (Teacher's Edition)*. 2007;(10):43-48+59.
- Liu M. Review and prospect of English corpus and corpus linguistics research in the last 40 years. *Journal of Shaanxi Institute of Education*. 2004;20(1):98-103.
- Liu M. On corpus-based teaching of English synonyms. *Journal of Shanxi Institute of Education*. 2006;22(4):98-103.
- Ma Y, Zhang C. Research on English reading vocabulary teaching with corpus assistance. *Chinese Journal of Education*. 2019;(S1):84-85+99.
- Pu JZ. Class linkage, collocation and word blocks in English vocabulary teaching. *Foreign Language Teaching and Research*. 2003;(6):438-445.
- Sun J. An experimental study of corpus indexing for high school English vocabulary teaching [master's thesis]. Shanghai Jiaotong University; 2007.
- Tan AH, Feng CY. A study on vocabulary teaching method based on corpus linguistics. *Journal of Southwest University for Nationalities (Humanities and Social Sciences)*. 2008;(S1):267-269.
- Tang JY, Ho AP. The use of corpora in foreign language teaching. *Foreign Language Teaching and Learning*. 2004;(05):42-45.
- Wang JY. Corpus-based English vocabulary teaching: Rationale and application. *Journal of Foreign Languages*. 2012;(04):127-130. doi:10.16263/j.cnki.23-1071/h.2012.04.029.
- Wang J, Ge G, Chun F, Peng R, Wang F, Q FL. The application of corpus in English vocabulary teaching. *Journal of the Fourth Military Medical University*. 2005;(S1):124-126.

36. Wei N. Corpus-based and corpus-driven word collocation research. *Contemporary Linguistics*. 2002;(02):101-114+157.
37. Wen Q, Wang J. A corpus of Chinese university students' English-Chinese-Chinese-English oral and written translation. Beijing: Foreign Language Teaching and Research Press; 2009.
38. Wu WW, Wang XX. Web-based corpus and its application in English teaching. *Research on Electrochemical Education*. 2002;(10):54-57. doi:10.13811/j.cnki.eer.2002.10.012.
39. Wu J. An empirical study of corpus-based vocabulary teaching in high school English [master's thesis]. Northwest Normal University; c2013.
40. Hsieh JC. The use of personal English teaching corpus. *Foreign Language Teaching and Learning*. 2004;(02):34-36.
41. Xie JC. Vocabulary corpus analysis of secondary English textbooks. *Theory and Practice of Foreign Language Teaching*. 2010;(01):55-61.
42. Xu JJ. The communicative view of vocabulary-centered pedagogy-theoretical traceability and reflection. *China Foreign Language Education*. 2009;(4):38-45.
43. Xiang Z. Corpus-based design of university English vocabulary hierarchical teaching. *Journal of Qiqihar University (Philosophy and Social Science Edition)*. 2015;(12):173-174+177. doi.
44. Xu M, He A. Schema theory, corpus linguistics and foreign language teaching. *Journal of the PLA College of Foreign Languages*. 2004;(06):45-49.
45. Xu M, Lee J. Research on the teaching of English vocabulary in China: A review and prospect. *Foreign Language Society*. 2007;(03):69-74+96.
46. Xu Q. A new tool for teaching English vocabulary: the use of online corpora. *Global Education Perspectives*. 2009;(08):90-93.
47. Xu Q. A vocabulary-centered teaching method based on online English corpus. *Shanghai Educational Research*. 2010;(01):89-90. Doi:10.16194/j.cnki.31-1059/g4.2010.01.004.
48. Yan C. A study of corpus-based vocabulary learning in English classroom teaching. *Chinese Journal of Education*. 2018;(S1):94-96.
49. Yu Y. Data-driven vocabulary teaching - An exploration of computer- and corpus-based research teaching. *Foreign Language e-Learning*. 2009;(02):58-62.
50. Zhang YL, He A. Corpus-based vocabulary learning in high school English. *Foreign Language Teaching in Elementary and Middle Schools (Middle School)*. 2006;(08):7-12.
51. Zhao X. A corpus-based study of lexical styles. *Foreign Language e-Learning*. 2010;(04):27-31.
52. Zhen F. Corpus Data-Driven Foreign Language Learning-Ideas, Methods and Techniques. *Foreign language community*. 2005;(4):19-40.
53. Chu V. Data-driven learning: A new trend in English vocabulary teaching. *Foreign Language e-Learning*. 2011;(01):46-50+59.