THE ROLE OF CORPUS LINGUISTICS IN LEARNING VOCABULARY

Kakhorov Maksud Usmon o'g'li

an assistant of the department of

Linguistic support of intercultural communication

Uzbekistan State World Languages University

email: maqsudqahhorvov19@gmail.com

Abstract: This article examines the methods of teaching vocabulary using a corpus-based approach and also provides some data on how to employ a corpus-based approach in a teaching environment. Furthermore, it reveals some information about the reasons for using a corpus to teach vocabulary and its benefits.

Keywords: Frequency, corpora, adult fiction, high-frequency words, word families, academic words, written language.

A corpus is as a collection of natural language examples consisting of some sentences in a set of written texts or records collected for linguistic study, and the texts are then arranged systematically[1]. Corpus is declared "natural" since the texts collected are those produced and used naturally and as is or not made up. The texts include textbook journals, novels, newspapers, magazines, and records of conversation broadcasts, interview results, and many more. Corpus linguistics is a complete system containing methods and principles to apply corpus in linguistic research and teaching

---

[1] Pishghadam, R., & Zabihi, R. (2012). Crossing the threshold of Iranian TEFL. Applied Research on English Language, 1(1), 57-71. https://doi.org/10.22108/are.2012. 15446
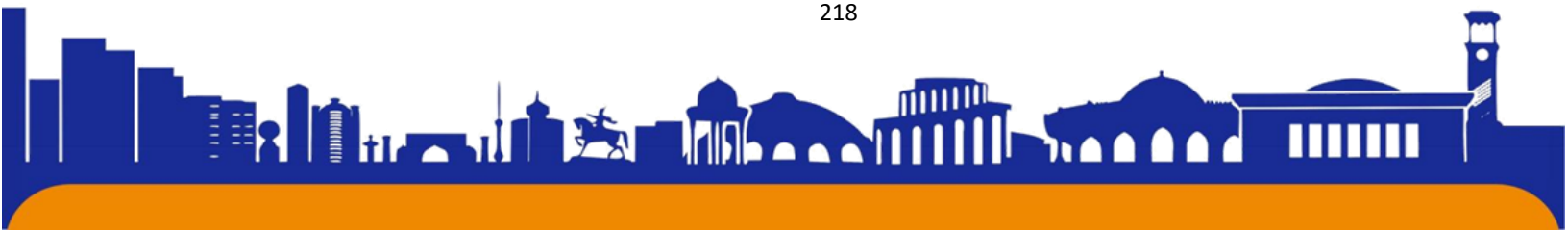
or learning[2]. Corpus linguistics is also a field focusing on a set of procedures or methods for learning a language [3]. Based on these definitions, we may conclude that corpus linguistics is linguistic research that uses examples of daily or natural language stored in a corpus. Corpus linguistics is suitable for application in this research since corpus linguistics has the features needed to achieve the research aim. The linguistic features include frequency in corpus linguistics, referring to the appearance of a word in a corpus or text[4]. Not only used to count the appearances of a single word, the frequency also allows the counting of grammatical, semantic, or other categorical frequencies. Frequency can also guide the researcher to wider findings. Frequency in corpus linguistics shows how many times a word appears in a corpus.

Frequency analysis allows the researcher to recognize words often appearing in certain situations in real life.

These frequencies used in any language corpus can be utilized to teach vocabularies for language learners due to the fact that there can be found authentic material, meaning words that are common in those languages. Furthermore, According to Varley technology and computers have become one of the main aspects of human life. No one can deny the fact that technology has a great effect on the way people lead their lives. However, according to Breyer[5], teaching is one of the areas where technology has not had a strong impact. Corpus linguistics is one of the technology-based tools that could be very useful in teaching, but it has still not been widely used or tested. Nevertheless, in the last 30 years, the use of corpora in classrooms has started to develop. Another feature of a corpus is that it is a principled collection of texts

---

[2] Callies, M. (2019). Integrating corpus literacy into language teacher education. Learner Corpora and Language Teaching, 92, 245-263. https://doi.org/10.1075/scl.92. 12cal

[3] Gholaminejad, R., & Sarab, M. R. A. (2020). Academic vocabulary and collocations used in language teaching and applied linguistics textbooks: A corpus-based approach. Terminology, 26(1), 82-107. https://doi.org/10.1075/term.00043.gho

[4] Crossley, S. A. (2020). Linguistic features in writing quality and development: An overview. Journal of Writing Research, 11(3), 415-443. https://doi.org/10.17239/ jowr-2020.11.03.01

[5] Breyer, Y. (2008). Learning and teaching with corpora: Reflections by student teachers. Computer Assisted Language Learning, 22(2), 153-172.

available for qualitative and quantitative analysis. This definition is useful because it captures a number of important issues. On the other hand even the corpus based tools has been developed, L2 classes has no opportunity to provide internet access classes. So Students can use teacher corpora, classroom corpora. It is, of course, natural that learners and teachers of any languages may pose questions about the use of corpora in learning languages. The answer is apparent and it is the exploitation of texts to pick up the vocabulary.

The collection of text or corpus is called as corpora. A concordance lets readers see texts in context. Some tools can be used to analyze corpora, for example, AntCont, Concordance, ClIC, and CorpKit. Various tools provide the creation of frequency information, such as the language frequency database which lists all words appearing in the corpus and defines how many times each unit happens in that corpus. In order to understand the vocabulary, Corpus-based research is commonly used and can be done quantitatively and qualitatively. Thompson and Sealey[6] have performed a corpus-based analysis to examine the children's literature vocabulary. They compared the children's literature corpus to a corpus of adult literature and newspaper articles to determine whether children's written language has different linguistic properties relative to the adult text. Furthermore, their finding shows that the vocabulary in children's literature shares many of the language's characteristics in adult fiction but, to a lesser extent, the news text's vocabulary profiles. Moreover, corpus-based research can examine almost any language pattern. It can analyze the lexical, structural, discourse, phonological, and morphological aspects of a language. Corpus-based research could enhance students' ability to identify useful phrases and common collocations, identifying the structure and nature of both spoken and written discourse. Atar and Erdem[7] also state that corpus-based research provides data for language studies. Corpus-based research includes holistic information about language structures because the data obtained from corpora

[6] Thompson, P., & Sealey, A. . Through children's eyes ? Corpus evidence of the features of children's literature. In International Journal of Corpus Linguistics (Vol. 12, Issue 1). John Benjamins Publishing Company, 2007

[7] Atar, C., & Erdem, C. . The advantages and disadvantages of corpus linguistics and conversation analysis in second lan guage studies. Proceedings of IX Scientific and Practical Internet Conference of Young Scientists and Students, November. https://www.researchgate.net/publication/337858444, 2019

is commonly analyzed by computer, and evidence is used concretely. Corpus-based research does not focus on single contexts. Multiple contexts were drawn to grasp the general understanding of the word's nature in the real-life data. Therefore, vocabulary can be easily mastered by using corpora. Vocabulary is a set of words that speakers of language use. Nation[8] states that high frequency words are words that are frequently used in language production. Those words are classified from the 2,000 most frequent word families—the words used frequently in a formal or informal situation. The words occur in the written and spoken text, such as newspapers, conversations, novels, and academic texts. Nation (2008) also suggests that these words should be considered as the target words in the vocabulary development program for the basic level. Academic words are words that can be found in academic texts. Academic words are different from the list of 1,000 or 2,000 of high-frequency words. These words occur in the newspaper, children's books, very formal conversation, and academic writing. These words consist of 570-word families and are commonly known as the Academic Word List (AWL). Nation[9] also adds that university students who use English for their academic tasks should learn Academic vocabulary. However, It is important that academic vocabulary is usually learned after the students mastered the high-frequency words. Berne and Blachowicz[10] state that vocabulary learning is an important part of English language learning. The learning of a new word in articles, books, or the internet sound very critical. It is also fundamental to language teaching and is of utmost importance to a language learner. More repetitions of vocabulary within a context could help a learner to learn and acquire vocabulary better. The use of interesting story, interesting articles, or news could help students to recognize the vocabulary. The

---

[8] Nation, I. S. Teaching and learning vocabulary. Newbury House,1990

[9] Nation, I. S. Teaching vocabulary: Strategies and techniques. Heinlee Cengage Learning, 2008

[10] Berne, J. I., & Blachowicz, C. L. Z. What reading teachers say about vocabulary instruction : Voices from the classroom. 62(4), 314–323. DOI:10.1598/RT.62.4.4, 2008
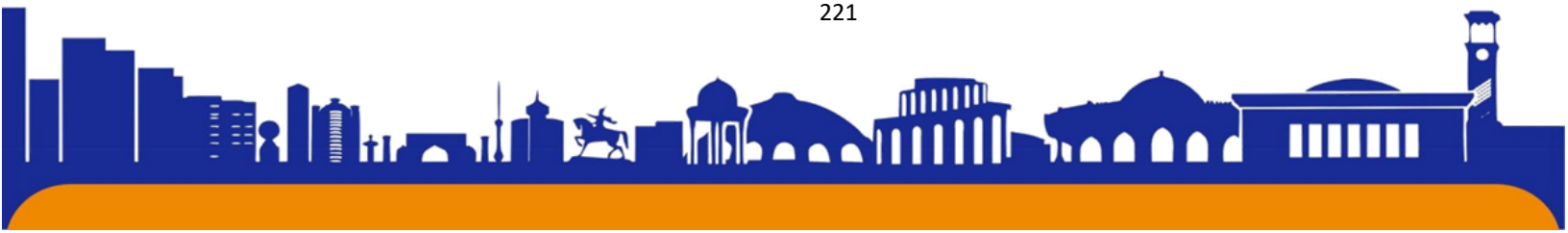
incidental process of vocabulary learning would facilitate vocabulary mastery. Chi and Lip[11] found that the most widely used and effective vocabulary learning techniques are

1) repetitive spelling of the word in the mind,
2) analyzing the word by breaking down the fragments of sound,
3) recalling terms by doing a project,
4) asking classmates for the meaning of the term.

To sum up, simple contextual use and regular pronunciation of the word and its meaning are commonly used in any language. They best-used those strategies for learning new words. It can be inferred that memory plays an important role in determining the vocabulary mastery. The vocabulary learning techniques that are commonly done refers to the repetition of the word. Therefore, a crucial factor in successful vocabulary learning was putting new words into practice that they have only learned. Repetition could be done through corpus-based research because corpus-based research could expose learners with words from real data.

References:

1) Pishghadam, R., & Zabihi, R. Crossing the threshold of Iranian TEFL. Applied Research on English Language, 1(1), 57-71. https://doi.org/10.22108/are.2012. 15446, 2012

2) Callies, M. Integrating corpus literacy into language teacher education. Learner Corpora and Language Teaching, 92, 245-263. https://doi.org/10.1075/scl.92. 12cal, 2019

3) Gholaminejad, R., & Sarab, M. R. A. Academic vocabulary and collocations used in language teaching and applied linguistics textbooks: A corpus-based approach. Terminology, 26(1), 82-107.
https://doi.org/10.1075/term.00043.gho, 2020

---

[11] Chi, P., & Lip, H. Investigating the most frequently used and most useful vocabulary language learning strategies among Chinese EFL postsecondary students in Hong Kong. Electronic Journal of Foreign Language Teaching, 6(1), 77–87. http://e-flt.nus.edu.sg/, 2009

4) Crossley, S. A. Linguistic features in writing quality and development: An overview. Journal of Writing Research, 11(3), 415-443. https://doi.org/10.17239/ jowr-2020.11.03.01, 2020

5) Varley, S. I'll just look that up in the concordance: Integrating corpus consultation into the language learning environment. Computer Assisted Language Learning, 22(2), 133-15, 2008

6) Breyer, Y. Learning and teaching with corpora: Reflections by student teachers. Computer Assisted Language Learning, 22(2), 153-172, 2008

7) Thompson, P., & Sealey, A. . Through children's eyes ? Corpus evidence of the features of children's literature. In International Journal of Corpus Linguistics (Vol. 12, Issue 1). John Benjamins Publishing Company, 2007

8) Atar, C., & Erdem, C. . The advantages and disadvantages of corpus linguistics an d conversation analysis in second language studies. Proceedings of IX Scientific and P ractical Internet Conference of Young Scientists and Students, November. https://www.researchgate.net/publication/337858444, 2019

9) Nation, I. S. Teaching and learning vocabulary. Newbury House,1990

10) Nation, I. S. Teaching vocabulary: Strategies and techniques. Heinlee Cengage Learning, 2008

11) Berne, J. I., & Blachowicz, C. L. Z. What reading teachers say about vocabulary instruction : Voices from the classroom. 62(4), 314–323. DOI:10.1598/RT.62.4.4, 2008

12) Chi, P., & Lip, H. Investigating the most frequently used and most useful vocabulary language learning strategies among Chinese EFL postsecondary students in Hong Kong. Electronic Journal of Foreign Language Teaching, 6(1), 77–87. http://e-flt.nus.edu.sg/, 2009