# IMPACTS OF LANGUAGE CORPUS IN VOCABULARY DEVELOPMENT OF ENGLISH LANGUAGE LEARNERS

Toshpulatova Mehriniso Kilichevna
Doctoral Student, Termez State University
Phdtoshpulatovamehriniso@gmail.com

**Abstract**
The study investigates the beliefs of English language teachers on applying Language Corpus in vocabulary development of English Language learners Drawing on a substantial database of concordance in use of language corpus, students' assessments and analysis of corpus based worksheets have been practiced through software concordance. The first part of paper is a brief review of literature that presents some theoretical information on vocabulary development of ELLs and how effective using corpora in language classes as a means of in offering real authentic examples of language stressing lexical markers. The second part of this paper aims to suggest Data Driven Learning (DDL) in order to analyze learners' vocabulary learning process in authentic contexts whereas the last section of this research paper indicates how using language corpus affects the vocabulary development of L2 learners.

**Keywords:** Language corpus, concordance, vocabulary teaching, vocabulary learning, lexical phrases

## Introduction
With the development of advanced technology and innovative methods, teaching of second language acquisition has gained a new framework by means of "Corpora" an effective and powerful instrument in teaching vocabulary as a foreign language. Corpus analysis has a big potential for language learning with the multidimensional aspects. This analysis system easily provides practical and quick access to the nature of linguistic data as an innovative teaching technique that is not based on traditional tools.

Rather than outdated traditional techniques or approaches, effective verity of the innovative instructions need to be adopted and supported if we would like to transform our classroom into an effective language learning environment. Teaching may be thought as an easy and a simple task, yet to be able to provide accurate and durable knowledge to the world of a learner is more sophisticated. However, the suggestive question is whether to be able to teach a language in an effective way making use of the innovative methods which fulfill the necessities of the time or with outdated methods which hardly keeps up with dizzying changes of life. Therefore,

revolutionizing vocabulary teaching according to the necessities of the world based on the living language which finds a place in the corpus may play a crucial role.

Utilizing a corpus for teaching is a handy way in language classrooms with the operative stance towards the core of language acquisition, whether the role of teacher is equally significant, in terms of vulgarizing corpus data until students comprehend how to analyze these sources. The aim of this paper is to clarify to what extent the implementation of Corpora effective instrument in vocabulary development of 4th graders at A2 Low-Intermediate Level. Further, the major role of "Corpus Program" as a linguistic resource is to be able to improve either students' learning skills or strategies likewise by cultivating their attitudes in academic and linguistic fields; such as, reading, writing and speaking. It also investigates if the bounded correlations among the students for the same levels of proficiency in corpus-based activities really reveal its efficacy in a sharp and effective way or not.

Vocabulary teaching seems to had been neglected - to use a colloquial expression-was left to take care of itself -for years and was of the secondary importance after grammar due to the fact that grammar was thought to be the core of the language competence and the ultimate goal of language teaching was seen to be able to produce grammatically correct sentences by prominent methods of 1960s and 1970s.With the emergence of communicative methods, the focus of language teaching naturally shifted from being accurate to being intelligible. It is argued that knowing rules alone does not make any sense in terms of using language; in order to communicate, one first needs to know words, chunks and word patterns, in short lexical knowledge, and then the rules to put these words into meaningful order, that is, grammatical competence. Therefore, both are indispensable and integrative part of language competence but considering the importance, it would not be wrong to say that lexical knowledge comes first and lies at the center of the ability to communicate effectively. If we clearly sum up with Wilkins' as cited in Donesch-Jezo, words, "Without grammar very little can be conveyed, without vocabulary nothing can be conveyed", so each is a component element supporting each other at the same platform.

Learning vocabulary in the second language is a multidimensional process which requires gaining more than one aspect of knowledge about the word. Besides form and meaning, a L1 learner already possess the knowledge of L1 words; what they refer to, what they connote, with what words they are linked, however, the L2 learner has to set up a new lexical construction by himself or herself or with the help of an instructor through explicit or implicit teaching and thus, the need of vocabulary teaching appears. But how it should be done? For reasons beyond this scope of this discussion, we are not going to mention how vocabulary should be taught in detail but

the point which related to our questions will be developing vocabulary through Corpus-based activities. Also, the fact that students have tendency dealing with new technology while learning a language since they are surrounded with various kind of smart technology devices and these devices make them more motivated is unquestionable. Therefore, they are expected to learn better when they are exposed to different teaching methods integrating technology such as corpus and their language development will gain great acceleration.

With the diffusion of the computer in educational field, corpus-based approach to language teaching become more wide-spread as it makes finding, counting and displaying a large number of examples of the words quicker and easier; the number of the data in corpuses are increasing day by day since more and more texts are written electronically and programs are developed to assist the process. Postman (1992) states advanced computer technology makes the existence of innovative methods possible to conduct vocabulary or grammar studies within a unique scope of ease, so the corpus linguistics which relies on computer-assisted techniques to analyze tremendous databases of naturally occurring language has a great facilitative potential in vocabulary instruction.

The Brown Corpus which was developed by Kučera and Francis in 1960 can be regarded as the earliest version and touchstone of English-language corpus and it was followed by the London-Lund Corpus, in 1980 an initial example of a spoken corpus and the Collins Birmingham University International Language Database or COBUILD that was prepared with the aim of writing a dictionary in 1987. Since then corpora have been used in ELT and especially dictionaries have started to be written corpus-based.

Using corpora for vocabulary teaching offers learner a valuable source about the possible collocation of the words and its co –occurrence with the other words about how language is used in real-life situations. 'Corpus' which, can be described as a large compilation of written or recorded texts or as McCarten defines, ''massive collection of texts has that given us access to a wealth of information regarding spoken and written English…'' (p. 5). Its size can vary from a few words to millions of sentences and the form in which it provides linguistic knowledge is called as concordance. According to Donesch-Jezo "A concordance is a list of all the occurrences of a particular word or phrase in a corpus, presented within the context (usually a few words to the left and right of this word)" . A concordance can also descried as a kind of screen display of a word used by many speakers for different situations (McCarten 2007, p.7). Concordances allow learner to meet all examples of a word in different

contexts. Since the words make sense in the context they are used, they need to be given meaningful context while being taught.

## Corpus Based Vocabulary Knowledge

In other words, vocabulary in context helps learner get the meaning from connection with the situation and the other linguistic elements such as verbs, noun, adverbs etc. used around that word. The knowledge of a word means more than just knowing the meaning word in isolation and it also includes the knowledge of the words that it co-occurs, in other words, the words that it collocates with. The absence of collocation may cause deviant and odd utterances such as make homework, turn on the window, beautiful man etc.  John Sinclair (as cited in Kilgarrif, n.d.), leader of the COBUILD project, notes that defining  a language without  mentioning  corpora is improper and only by looking at a corpus  we can see a how a language functions. As Gardner states while students learn the morphological relationships between phrases or multiword matters, they also achieve to be to delve into the core of lexical words, homonymy and polysemy via computer based corpora, as a result, all students have a great chance to harmonize all these linguistic knowledge and practice the in a different language platform in a vigorous way in the methodology of modern language teaching. Let's look how corpora works with the following example of the verb *see* in Figure 1 in order to have a good grasp of the notion.
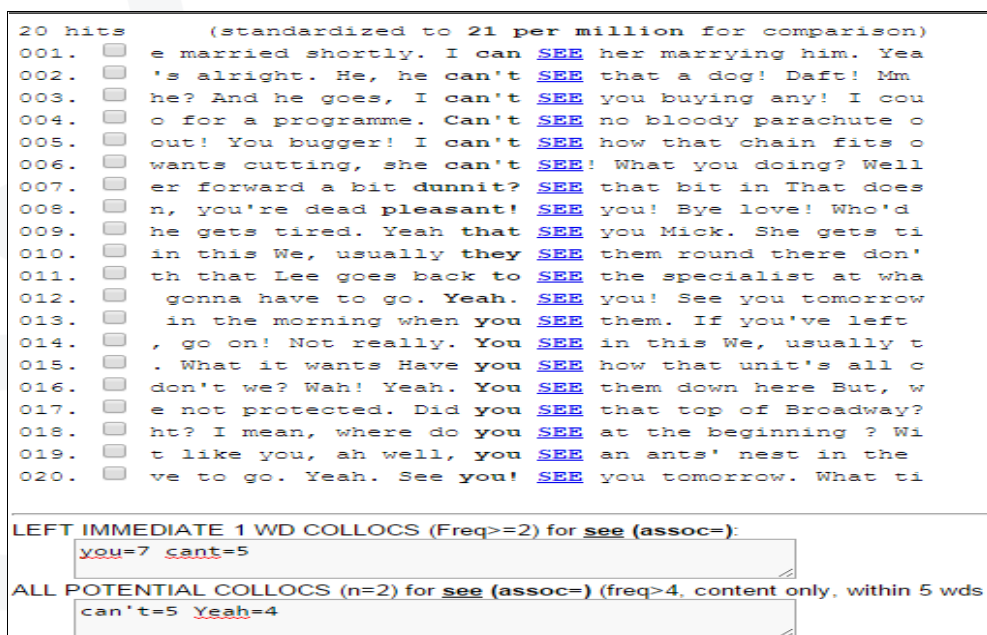


Figure 1: Concordances of the verb *SEE* from The British National Corpus Spoken (BNC)

Figure 1 shows a representative selection of verb ''see'' from the Corpus in which each example is given with a concordance. The word is put at the center of the page and written bold in some ways and preceded and followed by rest of the text. This is very helpful to decide which words go before or after a verb, such as objects or adverbs. It might prevent us listening to our intuitions which may cause any misuse. When we look at the calculation below the chart, let's say you: 7 and can't =5, we can easily reach the information about the frequency of each word and it is indisputable that knowing how frequent a word or a pattern facilitates language teachers to select the more frequent words or patterns and present them first and postpone the complex structures until another stage. Whereas frequency is not the determinant to decide the order, it can make teaching more effective. Therefore, corpora has also become a reference for teacher to derive the most frequent and by means of frequency list, decide what vocabulary to teach.

Another benefit of using corpora in language education is that it supplies more realistic examples of language usage including the nuances and complexities of natural language. Authentic usage of the language that is presented through corpora meets the learner with the actual usage. Maddalena emphasizes that the use of authentic and real-life examples with L2 learners is more useful than examples that are created by the teacher and do not simulate real use of language. It is also irrefutably accepted that being exposed to authentic language increases awareness of target culture and approximates the learner to natural usage of target language.

On the other hand, course books and materials written for created for foreign language teaching mostly seem to overlook this simple principle. To illustrate, some scholars have utilized from corpus data to review existing ELT syllabuses and teaching materials and they found out that the grammar written in textbooks for teaching English differs considerably from the daily usage of English. As Mindt observes (as cited in McEnery & Xiao, 2010, p. 13) that one of the most common problem with course books is that they offer a sort of English which cannot be heard out of the classroom and as a result, learners naturally have difficulty in establishing effective communication with native speakers. Considering these problems, writers such as McCarthy, McCartney and Sandiford published Touchstone course book series which is an innovative book based on research into the Cambridge English Corpus and aims to teach English as it is actually used.

# References

1. Conrad, S. M. (1999). The importance of Corpus-Based Research for Language Teachers: Iowa State University
2. Dee, G. (2006). Validating the Construct of Word in Applied Corpus-based Vocabulary Research: A Critical Survey. Oxford University Pres.
3. Davies, M. (2004) BYU-BNC. Based on the British National Corpus from Oxford University Press. Available from http://corpus.byu.edu/bnc/.
4. Donesch-Jezo, E. (2013). Using Language Corpus in Teaching Foreign Language Vocabulary. International Multidisciplinary e-Journal. 2 ( 1 ), 11-25. Retrieved from http://www.shreeprakashan.com/Documents/
5. Kilgarriff, A. (n.d.). Corpora in the classroom without scaring the students. Retrieved from http://www.kilgarriff.co.uk/publications.htm
6. Maddalena, S.R. (2001). An investigation into how corpus analysis may be used in the second language classroom to solve some of the problems surrounding non-native speakers' understanding of seemingly synonymous words. Retrieved from ERIC database (ED458795).
7. McCarten, J. (2007). Teaching vocabulary: lessons from the corpus, lessons for the Retrieved from Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the and beyond. In S. Braun, K. Kohn & J. Mukherjee (Eds.), Corpus Technology and Language Pedagogy (pp. 5-24). Frankfurt am Main: Peter Lang.
8. Binkai, Jiao. (2012). An empirical study on corpus-driven English vocabulary learning in China. Canadian Center of Science and Education. Retrieved from http://dx.doi.org/10.5539/elt.v5n4p131