

Utilizing corpus linguistics to instruct English

Sunnatillo Baxtiyorovich Otajonov
University of Exact and Social Sciences

Abstract: The study explores the usage of electronic corpora in the context of English training. The methodological promise of corpus linguistics is duly acknowledged. Despite the undeniable effectiveness of using corpus data for teaching and learning a foreign language, English educators have not given sufficient attention to the topic of corpus linguistics. The corpus method can be considered a valuable tool for developing practical skills and abilities in foreign language education when organizing students' individual work. This is because it enables pupils to enhance their language proficiency via practical application. In order to effectively teach the theory and practice of the English language, it has been determined, through a review of existing literature and analysis of data from the British National Corpus, that it is necessary to develop specific guidelines for both teachers and students.

Keywords: case, corpus linguistics, communication, information

Using information technology, it is now feasible to study a foreign language at a completely different level than it was before conceivable. This is made possible by the progress of science nowadays. There is a correlation between the growth of higher education in the present day and the potential of the most recent teaching aids. The informatization of vocational education is one of the most essential prerequisites for enhancing the quality of vocational education currently available. The informatization of society is the first step in the formation of an information and communication environment. Within this environment, the primary resource is a person who is able to acquire, save, and work with information, as well as apply it creatively in life and in professional activities. Additionally, this person is able to participate in the processes of searching for and creating new knowledge. Knowledge is a vital source of intellectual, personal, and professional growth for individuals who are preparing to become specialists. The preparation of such a specialist can only be accomplished via the utilization of innovative methods that are centered on the individual's advanced development. "The capability of resolving standard problems of professional activity on the basis of information and bibliographic culture by utilizing information and communication technologies while taking into consideration the fundamental requirements of information security." This formulation makes it abundantly obvious that students are required to cultivate the skills and capabilities necessary to use the technologies that have been listed at a professional level.

Using information technology, it is now feasible to study a foreign language at a completely different level than it was before conceivable. This is made possible by the progress of science nowadays. There is a correlation between the growth of higher education in the present day and the potential of the most recent teaching aids. The informatization of vocational education is one of the most essential prerequisites for enhancing the quality of vocational education currently available. The informatization of society is the first step in the formation of an information and communication environment. Within this environment, the primary resource is a person who is able to acquire, save, and work with information, as well as apply it creatively in life and in professional activities. Additionally, this person is able to participate in the processes of searching for and creating new knowledge. Knowledge is a vital source of intellectual, personal, and professional growth for individuals who are preparing to become specialists. The preparation of such a specialist can only be accomplished via the utilization of innovative methods that are centered on the individual's advanced development. "The capability of resolving standard problems of professional activity on the basis of information and bibliographic culture by utilizing information and communication technologies while taking into consideration the fundamental requirements of information security." This formulation makes it abundantly obvious that students are required to cultivate the skills and capabilities necessary to use the technologies that have been listed at a professional level.

For the purpose of providing pupils with such preparation, instructors themselves need to possess the requisite abilities.

According to the current circumstances, it is not sufficient for a teacher to only be a user; it is required to discuss the possibility of expanding the teacher's competence in the field of information and communication technology, which is his professional characteristic and a component of pedagogical expertise. When discussing the two-level model of information and communication competency of a teacher that is utilized in pedagogical practice, E.A.Lahaina makes reference to the following points:

1) the degree of functional literacy, which refers to the condition of being prepared for action and includes the following: the possession of computer programs that are capable of processing text, numerical, visual, sound, and video information; the capacity to operate on the Internet and utilize its services; the ability to use equipment such as a scanner, printer, and other similar devices;

2) activity level (implemented activity) - utilize functional literacy in the sphere of information and communication technology in educational activities in an efficient and methodical manner in order to obtain excellent outcomes. As an additional point of interest, this model proposes that the activity level is subdivided into the implementation and creative sublevels. The first possibility involves the incorporation of specialized media resources into educational activities.

These resources are generated in line with the criteria for the content and methodology of a certain educational subject. The second possibility involves the development of our very own electronic instructional instruments. The activity level of the teacher's competence should be our objective for developing the aforementioned competence in pupils, since it is apparent that this is the most important aspect of the teacher's competence.

Let's investigate corpus linguistics, shall we? In both the theoretical and practical aspects of teaching a foreign language, the methodological apparatus of corpus linguistics is a productive instrument that has great potential. Corpus linguistics is a subfield of computational linguistics that focuses on the creation of general rules for the generation and utilization of linguistic corpora (text corpora) through the utilization of computer technology. A linguistic, or language, corpus of texts is regarded to be a vast collection of language data that is unified, structured, labeled, and philologically competent. It is delivered electronically and is meant to answer specific linguistic issues. Additionally, this phrase may also be used to refer to a system that is used for the management of textual and linguistic data. This system is known as a corpus manager (or corpus manager at various times). Specifically, the latter functions as a specialized search system that incorporates software for searching data inside the corpus in order to acquire statistical information and display the findings to the user in a style that is easy to use. As a result, the resources that are contained within linguistic corpora may be regarded as an essential component of the methodology that applies to the teaching of a foreign language. Additionally, it appears that electronic text corpora, or more specifically, corpus analysis, ought to be taken into consideration as a means by which a student of philology might independently get access to information.

For students of a foreign language, there is a sort of educational corpus that can be interpreted as an electronic corpus of texts. The primary objective of this type of corpus is to evaluate the texts in order to determine the techniques and efficacy of studying the language that is being studied. In the field of linguistic analysis, corpora of this kind may be utilized to discover lexical or syntactic faults that occur throughout the process of studying a foreign language. This helps to establish the prevalence of particular types of language defects and the settings in which they occur. For the purpose of generating strategies and methodological procedures for future correction in language education, such data is required. Additionally, these corpora are utilized in the process of teaching.

One of the ways in which we have gained expertise with corpora is by our participation in a class that focuses on the lexicology of modern English. The British National Corpus, the Corpus of Contemporary American English, and other similar corpora are only a few examples of the many distinct English-language corpora that are available currently. As part of the student's independent work, being familiar with

these corpora reveals a wide variety of options, not just for the researcher but also for the student.

The British and American forms of a word are often the focus of attention in a lexicology class. With the help of the British National Corpus that is already in existence, it is possible to carry out a comparative analysis. This analysis can compare not only the lexical and grammatical characteristics of the British National Corpus with the data from the Corpus of Contemporary American English, but also the frequency of use of particular words and the combinability capabilities of those words. There is also the possibility of taking into consideration the morphological and derivational structures of words. Therefore, in response to a query on the morpheme -ism based on these two corpora, what is returned are the first one hundred frequency words that include the morpheme that was given in the word structure. The following terms are included in the selection:

1: Corpus of Contemporary American English, which is data from the American corpus;

2: British National Corpus, which are the identical words but are found in the British corpus;

PM 1 is the frequency with which a word appears in Corpus of Contemporary American English for every one million instances of that term; PM 2 is the frequency with which a word appears in British National Corpus for every one million instances. Ratio is a percentage ratio expressed as a frequency.

As an illustration, take into consideration the word challenge, which is a phenomenon that may be related to the extensive vocabulary of contemporary English. The results of the query are returned in 0.930 seconds, and the historical corpus contains the following information about this word: the total number of instances of this word's usage over the historical era is 1120, of which 22 instances are stated in the year 1810, 392 instances in the year 1910, and 2031 instances in the year 2000 respectively. Additionally, the context in which this term is used is intriguing from the perspective of both the diachronic changes that have occurred in its semantics and the combinability abilities that it possesses, both grammatical and lexical. Additionally, the genre distribution of the word is something that should be taken into special attention. By utilizing the "keyword in context" function, one is able to acquire all of this information for the purpose of conducting a more comprehensive analysis.

It is common knowledge that in addition to monolingual corpora, there are also bilingual and multilingual corpora, each of which possesses certain pedagogic qualities and methodological functions that are unique to itself. The following are examples of didactic properties:

- a) the ability to speak several languages;
- b) the contextuality of search results; and

c) the capability of restricting the range of words that can be used.

Not only are parallel corpora helpful for improving speaking skills, but they are also helpful for building translation skills and talents. Because of this, the National Corpus of the Russian Language that is now in existence also includes parallel corpora, which include English-Russian as well.

A cursory examination of the capabilities of corpora substantiates the significance of utilizing them not only for the purpose of carrying out a variety of scientific research projects in and of themselves, but also for the purpose of attracting resources of this sort in the field of teaching foreign languages. It is only necessary to mention that ongoing efforts are being made to enhance and broaden the capabilities of existing corpora of many languages, including the corpus of the Russian language, which is known as the National Corpus of the Russian Language.

References

1. Sadovnikova O. E. Pryamoe i kosvennoe ispol'zovanie korpusov tekstov v zarubezhnoi lingvodidaktike [Direct and Indirect Use of Corpora in Foreign Linguodidactics]. *Magister Dixit*. 2013. No. 2
2. Zakharov V. P. *Korpusnaya lingvistika* [Corpus linguistics]. Petersburg, 2005. 48 pp.
3. Li Weina, Application of Case Grammar Theory in Basic English Teaching. *Journal of Guangxi Normal University*, 2011.
4. Fillmore, Charles J. The Case for Case. In Emmon Bach and Robert Harms. *Universals in Linguistic Theory*. New York: Holt - Rinehart - Winston, 1968.