

2. Романюк С. К., Денисюк Л. В., Данілова Н. Р. Формування комунікативної компетентності здобувачів-маркетологів під час вивчення англійської мови. *Наукові записки Національного університету «Острозька академія»: серія «Філологія»*. Острог : Вид-во НаУОА, 2022. Вип. 13(81). С. 104–107.

3. Тимощук Н.М. Комунікативний підхід до викладання іноземних мов у вищих аграрних закладах освіти. *Вісник Житомирського державного університету*. 2009. Випуск 43. Педагогічні науки. С. 92-95.

4. European Centre for the Development of Vocational Training. URL: The European Indicator of Language Competence | CEDEFOP (europa.eu)

5. Tymoshchuk N. The Implementation of E-Learning to Support Learning and Teaching Foreign Language: a Case Study in Ukrainian Higher Education. *Arab World English Journal (AWEJ)*. 2nd Special Issue on Covid 19 Challenges (2). 2022. 233-244. DOI: <https://dx.doi.org/10.24093/awej/covid2.1>

6. Zahorodna O., Saienko V., Tolchieva H., Tymoshchuk N., Kulinich T., Shvets N. Developing Communicative Professional Competence in Future Economic Specialists in the Conditions of Postmodernism. *Postmodern Openings*. 2022. 13(2). P. 77-96. DOI: <https://doi.org/10.18662/po/13.2/444>

УДК 378.147:81'243(045)

Tkachuck T. I.,

Candidate of Philological Sciences

ORCID: <https://orcid.org/0000-0002-4223-269X>

Matsera O. A.

ORCID: <https://orcid.org/0000-0001-7102-4497>

Vinnitsia Institute of Trade and Economics of SUTE

USING CORPORA IN THE LANGUAGE LEARNING PROCESS

The article examines the historical development of corpus linguistics. It has been determined that corpus linguistics arose in the 1960s due to the rapid growth of computer technology and has gone through three stages of formation and development.

The article reviews the advantages of using the corpus approach in teaching foreign languages. Analyzing authentic language materials from corpora helps students gain an in-depth understanding of real-life language situations and the use of language in context. Using corpus technologies helps enrich vocabulary and develop skills in text analysis, and reproduce the structure of language in appropriate communication situations. During training, there are opportunities to use explicit and implicit corpus-based approaches to develop lexical, grammatical, and phonetic skills. Today, corpora of written and spoken texts are successfully used in foreign language teaching and linguistic pedagogy.

Key words: *corpus technologies, corpus linguistics, concordance, foreign language teaching, authentic texts, explicit and implicit corpus-based approaches.*

The issue under discussion. In the realities of modern society, the flow of information is constantly increasing, so it is becoming increasingly important to automatically process language information. Recently, computers have played a meaningful role in the processing, storage, and selection of linguistic documents. This has served as motivation for the development of a relatively new field, corpus linguistics. The use of corpora is becoming extremely popular in language teaching and learning. Teachers and learners are applying to corpora for rapid, multi-level search and analysis of a wide variety of linguistic information, and concordances are recognized as a convenient tool for effective language learning. The use of corpora in teaching a foreign language in higher education institutions provides great opportunities for effective learning of the main functional features of language units in specific contexts.

Analysis of current research and publications on the topic. Recently, corpus technologies have been actively used in foreign and domestic linguistics. S. Buk, I. Bublyk, V. Kolkunova, V. Zhukovska, M. McCarthy, P. Baker, G. Leach, and others have made a significant contribution to the development of corpus technologies. Given that case technologies have been actively used abroad for a long time, their application in our country has only recently begun.

The study aims to review the historical path of corpus linguistics and consider ways of applying corpus technologies in the training of philology students.

Basic material presentation. Corpus linguistics is a relatively new linguistic field that began its active formation in the 1960s. This period was determined by the active development of computer technologies.

The researcher E. Tognini-Bonelli connects the history of corpus studies with the direct development of computer technologies and identifies the following stages:

The first stage (mid-60s–early 80s of the twentieth century): the period of acquiring knowledge about organizing and maintaining corpora of up to 1 million words, characterized by the lack of materials in electronic format and the need for manual typing.

The second stage (1980–2000) is divided into two periods:

a) The 1980s were marked by the emergence of scanners, when even primitive scanners could produce corpora of 20 million word occurrences;

b) The 1990s were marked by the expansion of computer typesetting capabilities, which facilitated access to large textual materials in electronic format and contributed to a significant increase in the size of corpora.

The third stage (from the beginning of 2000 to the present) is the period of electronic (virtual) texts that have never had a material form, providing enormous opportunities for creating corpora of any size [11].

In the United States, the period of formation of the foundations of corpus linguistics (the 1960s of the twentieth century) was marked by the flowering of transformational generative grammar. However, N. Chomsky disapproved of the corpus studies of that time, calling the corpus method of accumulating linguistic data inadequate and false for describing the generative capacity of natural language.

Other representatives of transformational generative grammar also showed no interest in the possibility of studying data on the real function of language.

As a result, the leading linguists and theorists of the time did not support or show interest in the new corpora. In the 1960s and 1980s, corpus linguistics was generally pushed to the periphery of the US language research paradigm [10].

In contrast to the USA, there were strong empirical tendencies in English linguistics, which implied the use of specific examples to test a linguistic hypothesis.

This has led to the emergence of many corpora and the development of corpus-based research on the English language. Corpus linguistics, along with other branches of linguistics that use computer technology as a technological tool, directly impacts traditional ways of collecting and storing language material—manual processing of written texts, dictionary indexes, and other forms [5].

As the name implies, corpus linguistics is based on the use of a corpus, i.e., a large amount of «live, real» language material extracted from a variety of sources and compiled into a computer system so that researchers, especially lexicographers, can study the meanings and linguistic patterns that emerge. From this point of view, corpus creation represents a new strategy not only as a method that emphasizes the significance of real facts of language but also as a theory that allows us to take a fresh look at the nature of previously formulated positions. The main task of corpus linguistics is to study language reality and language use in the process of communication and to determine and demonstrate its various aspects most correctly and accurately [2].

Corpora can be organized in different ways depending on the pragmatic goal of their creators, although they are also used for purposes not envisioned by the authors. They can be divided into monolingual and multilingual, general and specialized, parallel and comparable (translation corpora), synchronous and diachronic, etc. Texts, as constituent elements of a corpus, can be complete original verbal works or their parts. The materials collected in the texts are samples of the language used in the communication activities of the community, covering different genres and periods.

Today, most scientists believe that to train highly qualified and competitive specialists, it is necessary to constantly improve the forms and methods of teaching as well as rationally distribute theoretical and practical training [3].

The advantages of using corpus technologies include the following:

- modern corpus technologies allow for more effective formation of the professional foreign language competence of future specialists;
- the development of a corpus for a subject area ensures the adjustment and improvement of traditional methods of vocabulary selection in terms of making them more accurate and focused on the key concepts and lexical items of a given subject area;

- through the use of corpus technologies, pedagogical practice is enriched with modern science-based methods of selecting learning content that corresponds to the current state of basic science methodology;

- taking into account the results of corpus-based research contributes to the creation of an optimal learning model based on a corpus of lexical units and a set of exercises that meet the principles of professional orientation, representativeness, authenticity, and lexical richness;

- the use of the linguodidactic potential of the subject area corpus has a positive effect on the process of learning lexical units selected according to the needs of students;

- corpus-based research and the subject area corpus can serve as a basis for developing variable tasks that contribute to the more successful development of key competencies in students' learning [2].

In today's world, corpora are widely used in teaching and learning foreign languages. Teachers and students are increasingly using corpora for quick multi-level search and analysis of a wide variety of linguistic information, and concordances are recognized as a convenient tool for effective language learning [3]. A concordance program is a special text processing program that is assigned the linguistic task of finding a particular morpheme, word, or phrase in a context. For example, in the case of English, it is to find in a group of texts the options for using indefinite articles or all words ending in –«ing». As a result, the concordance program will display all the words with this ending along with the context, usually a line of text.

Currently, concordances are a modern and effective tool for text analysis that should be actively used in language teaching practice and linguistic tasks.

There are three ways to use corpora in teaching and learning foreign languages:

- 1) teaching materials based on corpus data,
- 2) illustrative examples from corpora,
- 3) exercises and tasks using corpus data (i.e., data-driven learning).

Recently, the use of corpus-based learning in the training of philology students in higher education institutions has been carried out in various ways, namely: use in

teaching various subjects; use in the preparation of teaching materials, textbooks, and manuals; use in lexicographic practice; use in the educational process in performing various tasks developed based on corpora (vocabulary, grammar, listening, reading, and writing); use for assignments [12].

There are two ways of using corpora for the study of foreign languages. Teachers can either encourage students to make use of the relevant software through hands-on, practical activities (explicit) or base their tasks on corpus data (implicit). In either case, corpus-based activities should be motivating, enlightening and make sense to learners.

Aston G. considers the possibility of using corpora to improve students' phonological competence. Students can be asked to compare the voiced lines of the concordance with their speech. Thanks to a transcript, students can perform the following tasks:

- a) listen to segments and then repeat them;
- b) repeat the sentences as they are spoken by the speaker without waiting for them to finish;
- c) read the segment aloud before listening.

In each of these tasks, students can record their version and then compare it with the corpus segment [6].

Babych S. suggests using corpora for reading and writing skills in the online environment. For example, the software can highlight all the link words in a text, which students can classify according to their function. These can be all the comparisons, points of view, and tense words. Students can then check themselves and correct themselves by moving the cursor over each word to see a box that indicates the semantic category. Students can also identify the topic sentence of each paragraph or discuss the identified passages. Another task could be to find arguments for and against the issue addressed in the text [7].

For students, the advantage of using corpora for learning is that they have access to authentic texts that have been taken from a variety of sources. From these texts, a huge number of examples can be collected, such as a word, a word usage, a

grammatical rule, a syntactic rule, etc. For students, the use of corpora and concordances is a unique opportunity to develop skills in analyzing information and drawing conclusions based on it to become researchers [1]. However, researchers in the field of corpus linguistics and methodologists note that the disadvantages of introducing corpus work are that it takes time to compile tasks from the large number of examples provided by the concordance to make them appropriate to the level of students and that neither teachers nor students want to move away from traditional teaching methods. At the same time, it would be useful to gradually introduce corpus-based assignments and invite teachers to professional development courses and training on the implementation of new teaching methods. Corpus-based tasks have been successfully implemented in foreign language textbooks for several years now, but they do not necessarily require students to use concordance and formulate their own rules.

Conclusions and further investigation prospects. As a result of the study, it can be noted that the use of corpus linguistics has a significant impact on the effectiveness of foreign language teaching. This approach makes it possible to bring the learning process closer to real-life language situations, providing an opportunity to learn a language in its natural context. Using corpora of texts allows observing the use of words, expressions, and grammatical structures in their natural environment. Analyzing authentic texts and observing the use of language in real-life communications helps to master not only grammatical structures and lexical items but also the sense of language nuance.

Thus, the use of corpus linguistics in foreign language learning is a very appropriate and promising approach that helps to gain a deeper and more practical understanding of the language.

References:

1. Aliieva A. A., Ostapchuk I. I. Korpusna linhvistyka u navchanni inozemnym movam. *Aktualni problemy filolohii ta perekladoznavstva*. 2023. Vyp.27. S.5-9.
2. Barkasi V.V., Kazakova M.M. Zastosuvannia korpusnoho pidkhodu pry navchanni inozemnykh mov u VZO. *Molodyi vchenyi*. № 4.4 (5.4). 2019. S. 53-56.

3. Zhukovska V. V. Vstup do korpusnoi linhvistyky: navchalnyi posibnyk. Zhytomyr: Vyd-vo ZhDU im. I. Franka, 2013. 142s.
4. Savchenko O.O., Stupnytska H.I. Korpusna linhvistyka v protsesi pidhotovky vchyteliv anhliiskoi movy. *Visnyk nauky ta osvity*. 2023. Vyp. 5(11). S.234-245.
5. Selivanova O. O. Korpusna linhvistyka. Suchasna linhvistyka: napriamy ta problemy: pidruchnyk. Poltava: Dovkillia-K, 2008. 712 s.
6. Tatsenko N. V. Metody korpusnoi linhvistyky v pidhotovtsi fakhivtsivfilolohiv. URL: https://essuir.sumdu.edu.ua/bitstream/download/123456789/81070/1/Tatsenko_corpus.pdf
7. Aston G. Learning phraseology from speech corpora. *Multiple affordances of language corpora for data-driven learning*. 2015. Vol.69. P. 65-84.
8. Babych S. Textual cohesion patterns for developing reading skills. *Multiple affordances of language corpora for data-driven learning*. 2015. Vol.69. P. 155-182.
9. Bennett G. R. Using corpora in the language learning classroom: Corpus linguistics for teachers. University of Michigan Press, 2010. 144 p.
10. Finegan E. Language: its structure and use. New York: Harcourt Brace College Publishers, 2004. 575 p.
11. McEnery T., Wilson A. Corpus linguistics: an introduction. 2nd ed. Edinburgh: Edinburgh University Press, 2001. 235 p.
12. Tognini-Bonelli E. Theoretical overview of the evolution of corpus linguistics. *The Routledge Handbook of Corpus Linguistics*. Routledge, 2010. 681 p.
13. Wichman A. Corpora and spoken discourse. *Language and computer studies in practical linguistics*. 2007. Vol. 62(1). P.73-74.