THE IMPORTANCE OF CORPUS LINGUISTICS

Fayzieva M. Ch.,

Teacher of English Linguistics Department,

Foreign Languages Faculty,

Bukhara State University, Bukhara

e-mail: m.ch.fayziyeva@buxdu.uz

Azimova N. G.,

Student of the group 11-7 Eng. 20,

Foreign Languages Faculty,

Bukhara State University, Bukhara

e-mail: azimovanigina2301@gmail.com

ABSTRACT

This article is about the importance and types of corpus linguistics. One of the most rapidly expanding approaches in modern linguistics is corpus linguistics. In the field of applied linguistics, large collections of written and spoken texts that have been specifically chosen to reflect certain linguistic contexts, such as casual conversation or academic writing, are analyzed.

Keywords: corpus, monolingual, parallel corpus, comparable, diachronic and synchronic corpus, specialized, multimedia.

INTRODUCTION

The study of corpus linguistics was established in the 1960s as a subfield of computational linguistics. Text corpora that have been built using computer technology are being used to generate theoretical frameworks. The term "linguistic (lexical) corpus" refers to a collection of tagged electronic texts and a sizable electronic database that are specifically chosen for solving linguistic problems. Using computer technologies, corpus linguistics investigates, analyzes, and studies language.

METHODS

In order to study how words, phrases, and language are used generally, a text corpus is an immense collection of text (typically many billion words) created by genuine language users. It is utilized in many different domains, including linguistics, lexicography, social science, the humanities, and natural language processing. A

corpus may also be used to create varied language databases that are used in the creation of software, including machine translation, predictive keyboards, spell check, grammar check, text-to-speech modules, text-to-speech comprehension systems, and many more. For scientific research in various branches of linguistics, corpus linguistics is also extremely useful. Corpus linguistics is a rapidly developing field in the world today. It is created by experts in corpus linguistics for the purpose of scientific research and language teaching. Educational corpus is equally important in learning mother tongue and foreign language. In teaching the language, the corpus is beneficial in showing an array of examples to show the size of the vocabulary and explain the possibility of using a word through one or another grammatical construction. Only a corpus has the chance and feature to reflect this constant updating of examples, which is important for language learning.

RESULTS

It is vital to note that corpus is impossible to effectively categorize into a certain category. Instead, corpora may include characteristics or attributes that may be used to organize them. One or more of these characteristics may exist in the same corpus. The most frequent type of corpus is a monolingual corpus. In only one language, it holds texts. The corpus is typically annotated for parts of speech and is used by a wide range of users for a variety of tasks, from highly practical ones, like verifying the proper usage of a word or looking up the most natural word combinations, to scientific ones, like identifying frequent patterns or new trends in language. Two or more monolingual corpora include a parallel corpus. The corpora are considered to be each other's translations. Fiction and its translation, for instance.

ANALYSIS

It is critical to mention that Orzubek Anarbayev follows the parallel corpus and gives an idea: "Parallel corpus is an important reality for the current era in which intercultural communication is widespread. Through parallel corpora in different language environments, universals in cultures, and specific mental characteristics of languages, it will be possible to identify real and lacunar units. A corpus of parallel texts also serves for the development of automatic translation; computer lexicography provides this development. Concordant programs using a corpus of parallel texts will be developed, and it will be possible to create different specialized dictionaries". Aligning the two languages requires matching equivalent parts, which are often phrases or paragraphs. A comparable corpus is one corpus in a group of two or more monolingual corpora that were all constructed using the same methods and are normally each in a different language.

DISCUSSIONS

Despite the fact that the corpora are not translations of one another, the content is comparable, making it possible to compare the findings between them. A diachronic corpus is a corpus holding texts from several historical eras and is used to investigate the evolution or change in the language. A synchronic corpus, on the other hand, contains writings that were written at the same time. It captures the language at that specific time. A specialized corpus consists of texts that are restricted to one or more domains, subjects, subject areas, etc. This corpus is employed to examine the usage of the specialized language. A multimedia corpus is a collection of texts that have additional audio, video, or other multimedia elements. Other corpora may contain photographs of the text's original manuscript or printed version or films in which the corpus text is read aloud.

In conclusion, it can be stated that the task of the corpus is large and undeniably unattainable. Today, corpus has become a work tool that saves time and labor. Changes that occur in the language over the years to the book-form dictionaries require a lot of effort and money. Corpus is an opportunity that solves problems in a short time and at no cost.

REFERENCES:

1. Abduraxmonova, N. Z. "Linguistic support of the program for translating English texts into Uzbek (on the example of simple sentences): Doctor of Philosophy (PhD) il dis. aftoref." (2018). Abdurakhmonova N. The bases of automatic morphological analysis for machine translation. Izvestiya Kyrgyzskogo gosudarstvennogo tekhnicheskogo universiteta. 2016

2. Ataboev, N. B. (2020). Functional features of the English corpus (in the example of COCA). *PhD thes. in phil.(189 p.). Tashkent.*

3. Атабоев, Н. Б. (2020). Инглиз тили корпусининг функционал хусусиятлари (СОСА мисолида) дис... ф. ф. док.Choshovna, F. M. (2022, January). The notion of family loyalty in literary works. In *Integration Conference on Integration of Pragmalinguistics, Functional Translation Studies and Language Teaching Processes* (pp. 107-109).

4. Jon McHardy Sinclair, Tillarni o'qitishda korporatsiyani qanday ishlatish kerak, Jon Benjaminlar, 2004-y.

5. Файзиева, М. Ч., & Туробова, Г. Ю. (2022). ВАЖНОСТЬ ИНКЛЮЗИВНОГО ОБРАЗОВАНИЯ В УЗБЕКИСТАНЕ. *INNOVATION IN THE MODERN EDUCATION SYSTEM*, *3*(25), 49-52.

6. https://cyberleninka.ru/article/n/korpus-lingvistikasi

7. file:///C:/Users/user/Downloads/korpus-lingvistikasi-va-parallel-korpuslar-tavsifi%20(1).pdf

8. https://uz.eferrit.com/

9. https://www.sketchengine.eu/corpora-and-languages/corpus-types/