CORPUS STUDY: EXPLORING LANGUAGE THROUGH REAL-LIFE DATA

Scientific supervisor: **Abdullajanova H**
Student of JSPU: **Kamalova Shakhzoda**

Annotation *(Abstract): This article explores the field of corpus linguistics, which analyzes language through large collections of authentic texts called corpora. The study discusses what a corpus is, the methods used in corpus analysis, and the importance and applications of corpus study in linguistics, education, translation, and computational fields.*

Keywords: *Corpus linguistics, corpus study, authentic texts, concordance analysis, frequency analysis, collocations, language research, applied linguistics*

Introduction

Corpus study, also known as corpus linguistics, is a branch of linguistics that focuses on the analysis of language using large collections of real-life texts called corpora. These corpora may include written texts, spoken conversations, academic articles, newspapers, social media posts, and many other forms of communication. Corpus study helps researchers understand how language is actually used, rather than how it is traditionally described in grammar books.

What Is a Corpus?

A corpus is a systematically organized digital collection of authentic texts. It is designed to represent a particular language, dialect, or type of communication. Corpora can be general (covering everyday language) or specialized (such as academic English, legal language, or learner English). Modern corpora are usually stored electronically, allowing researchers to analyze thousands or even millions of words quickly and accurately.

Methods Used in Corpus Study

Corpus study relies on computer-based tools and software to examine language patterns. Common methods include frequency analysis, which shows how often certain words or phrases occur, and concordance analysis, which displays a word in its surrounding context. Researchers also study collocations (words that commonly occur together), grammatical patterns, and discourse features. These methods provide objective and statistical evidence about language use.

Importance of Corpus Study

Corpus study plays a vital role in modern linguistics because it is based on real data rather than intuition. It helps linguists identify language trends, changes over time, and differences between spoken and written language. In language teaching, corpus research is used to design better textbooks, dictionaries, and teaching materials. It also helps learners understand natural usage, common mistakes, and appropriate word combinations.

Applications of Corpus Study

Corpus study is widely used in various fields. In lexicography, it helps create accurate dictionaries. In education, it supports syllabus design and language assessment. In translation and computational linguistics, corpora assist in machine translation and speech recognition. Researchers also use corpus study to analyze social issues, media language, and cultural communication patterns.

Conclusion

Corpus study has transformed the way language is researched and taught. By analyzing large amounts of authentic data, it provides reliable insights into how language functions in real contexts. As technology continues to develop, corpus study will remain an essential tool for understanding language in a scientific and practical way.

REFERENCES / USED LITERATURE

1.   McEnery, T., & Hardie, A. (2012). Corpus Linguistics: Method, Theory and Practice. Cambridge University Press.

2.   Biber, D., Conrad, S., & Reppen, R. (1998). Corpus Linguistics: Investigating Language Structure and Use. Cambridge University Press.

3.   Kennedy, G. (1998). An Introduction to Corpus Linguistics. Longman.

4.   Sinclair, J. (2004). Trust the Text: Language, Corpus and Discourse. Routledge.

5.   O'Keeffe, A., McCarthy, M., & Carter, R. (2007). From Corpus to Classroom: Language Use and Language Teaching. Cambridge University Press.

6.   Stubbs, M. (2001). Words and Phrases: Corpus Studies of Lexical Semantics. Blackwell.