

Corpus Linguistics: Analyzing Language through Large-Scale Textual Data

Hamza Jamal Jassim and Prof. Dr. Jagdish Joshi
Gujarat University, Ahmedabad, Gujarat, India

Abstract: *This article presents an in-depth exploration of corpus linguistics as a methodology for analyzing language using large-scale textual data, or corpora. It provides a comprehensive overview of the principles, applications, and impact of corpus linguistics on language research. The significance of corpus linguistics is highlighted in understanding language variation, discourse analysis, language teaching, and sociolinguistic research. The article emphasizes the availability of extensive corpora and digital technologies that have facilitated detailed linguistic analysis. It discusses how corpus linguistics enables the examination of language variation across genres, registers, time periods, and social contexts. The relationship between corpus linguistics and discourse analysis is explored, showcasing the data-driven approach of corpus-based discourse analysis. The role of corpus linguistics in language teaching and learning is addressed, along with its applications in sociolinguistic research. Overall, corpus linguistics has transformed language analysis, providing researchers with powerful tools to unravel the complexities of language (McEnery & Wilson 2001).*

Keywords: corpus linguistics, language analysis, large-scale textual data, language variation, discourse analysis, language teaching, sociolinguistic research.

REFERENCES

- [1]. Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.
- [2]. McEnery, T., & Wilson, A. (2001). *Corpus Linguistics: An Introduction*. Edinburgh University Press.
- [3]. O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge University Press.
- [4]. Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Blackwell Publishers.
- [5]. Biber, D. (1988). *Variation across Speech and Writing*. Cambridge University Press.
- [6]. Biber, D., & Conrad, S. (2009). *Register, Genre, and Style*. Cambridge University Press.
- [7]. Flowerdew, L. (2005). *An Introduction to Corpus Linguistics for Language Teachers*. Routledge.
- [8]. Tagliamonte, S. (2006). *Analysing Sociolinguistic Variation*. Cambridge University Press.
- [9]. Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. Longman.