

BRIDGING THE RESEARCH–PRACTICE GAP IN CORPUS LINGUISTICS: CHALLENGES AND DIRECTIONS

Z.U. Kulmatov,

Institute of International School of Finance Technology and Science (ISFT)

z.kulmatov@isft.uz

Abstract. Offering insights into real language use that shape dictionaries, grammars, and teaching resources, corpus linguistics has transformed language study. Still, there is a notable difference between corpus research and its practical use in language courses. This paper investigates the issues causing this disparity, such as teachers' ignorance of corpus approaches, insufficient training, and infrastructural limitations. Emphasizing teacher-centered initiatives and methodical training courses, it also addresses ways to close this gap. Addressing these problems will help corpus linguistics to better support language pedagogy and enable teachers to include data-driven learning into their instructional methods.

Through large-scale examination of genuine texts, corpus linguistics has developed into a potent tool for grasping language use. Its uses range from language teaching tools to grammar development to lexicography. Though these developments have been made, the incorporation of corpus-based techniques into regular classroom practice still falls short. Many teachers lack the required training or tools to make good use of corpora, which causes a gap between research results and teaching methods (Mukherjee, 2004). This paper looks at the elements causing this disparity and offers practical ideas to increase the use of corpus linguistics in language teaching.

Challenges in Bridging the Gap

1. Limited Corpus Literacy Among Teachers

The absence of corpus literacy among teachers is a major barrier. Many teachers lack knowledge of corpus tools or techniques, which makes it challenging for them to use corpus-based knowledge in their own classes (Mukherjee, 2004; Friginal et al., 2025). For instance, although modern teaching materials sometimes include corpus results—such as the sequencing of irregular verbs—teachers might not understand these tools are corpus-informed (Mukherjee, 2004). Their capacity to use these tools properly is hampered by this ignorance.

Corpus literacy is not just knowing how to operate corpus software; it also includes reading the data it offers. To guide their instruction, teachers must be able to examine other outputs, concordances, and frequency lists. Lacking this basic understanding, they find it difficult to include corpora into their lesson plans (Friginal et al., 2025).

2. Infrastructural Barriers

The incorporation of corpus linguistics into education is made more difficult by technological issues. Digital corpora may be less useful in schools with restricted access to computers and internet connection (Pérez-Paredes, 2023). The intricacy of corpus interfaces may also scare teachers lacking technical knowledge (Friginal et al., 2025). For example, many teachers lack the degree of technical knowledge needed to use AntConc or Sketch Engine.

Furthermore, schools sometimes give other educational software priority over corpus tools even when technology is available, so marginalizing their use (Pérez-Paredes, 2023). This emphasis shows more general systematic problems in which technology budgets are distributed depending on perceived urgent needs rather than long-term educational advantages.

3. Time Constraints

The rigorous character of current courses leaves teachers little opportunity to investigate innovative approaches such data-driven learning (DDL). Many teachers cannot afford the time and effort needed to create corpus-based resources given their responsibilities (Pérez-Paredes, 2023; Friginal et al., 2025). Usually charged with maintaining classroom dynamics and meeting curriculum requirements, teachers have little time for professional growth or testing new teaching approaches.

4. Absence of Systematic Training Programs

Pre-service and in-service education courses mostly lack corpus-based teacher training. Research indicates that even when teachers get some training in corpus techniques, it usually emphasizes technical issues over pedagogical uses (Ma, 2025). Teachers find it difficult to include corpora into their instruction without thorough training covering both corpus literacy and classroom implementation techniques. Good training courses should not only cover the fundamentals of corpus analysis but also offer useful illustrations of how to include these ideas into lesson plans. Workshops where educators collaborate to create corpus-based activities suited to their particular teaching settings could be part of this (Ma, 2025).

Strategies for Bridging the Gap

1. Teacher-Centered Corpus Activities

According to Mukherjee (2004), using teacher-centered corpus activities before presenting learner-centered approaches is crucial. Training courses can increase teachers' confidence in using corpora by emphasizing their needs and biases. For example, teachers could participate in group projects evaluating current curricula or creating new teaching resources by means of corpus data analysis.

2. Systematic Training Programs

Ma (2025) has suggested a two-step training program for teachers that consists of basic training in corpus ideas followed by hands-on experience with corpus tools. To guarantee broad adoption, such initiatives should be included

into teacher education courses. This strategy guarantees that educators have the practical skills and theoretical knowledge required for efficient corpus use.

3. Simplification of Corpus Tools

These tools can be more available to teachers by simplifying corpus interfaces and offering user-friendly tools. Scaffolded worksheets or guided activities, for example, can assist teachers negotiate difficult corpora such COCA or AntConc (Friginal et al., 2025). Creating web-based platforms or mobile apps that simplify corpus analysis would also help to lower obstacles even more.

4. Collaboration Between Researchers and Educators

Researchers should work with educators to create feasible corpus-based exercises fit for classroom requirements. This cooperation guarantees that research results are converted into practical teaching tools (Mukherjee, 2004). Joint research on the efficacy of corpus-based techniques in actual educational environments or co-authoring teaching resources could be part of collaborative projects.

5. Advocacy for Resource Allocation

Educational policy makers have to set aside funds to assist the incorporation of corpora into classes. This covers money for professional development courses and technology upgrades (Pérez-Paredes, 2023). Advocacy campaigns should underline the long-term advantages of corpus-based instruction, including better student performance and strengthened teacher professional development.

Future Directions

1. Empirical Studies on Effectiveness

More empirical studies are required to assess the effect of corpus-based teaching strategies on student outcomes at all proficiency levels (Friginal et al., 2025). Such studies can offer evidence-based justification for using these strategies. Longitudinal studies, for instance, could evaluate how corpus-based instruction affects language acquisition over time in comparison to conventional approaches.

2. Development of Corpus-Based Materials

Developing age-appropriate, context-specific corpus-based materials will help to fill in deficiencies in present resources, especially for younger students or novices (Ma, 2025). Designed with teacher involvement, these resources should fit classroom requirements and be practical to use.

3. Integration with Second Language Acquisition Theory

Combining knowledge from second language acquisition theory with corpus linguistics can improve awareness of interlanguage development and guide teaching methods (Mukherjee, 2004). This combination could result in more complex language teaching strategies considering both linguistic patterns and student processes.

Though it presents difficulties, the research-practice gap in corpus linguistics also creates chances for development. This disparity can be closed by tackling problems including teacher training, availability of resources, and

cooperation between academics and teachers. Empowering teachers with the knowledge and tools to use corpora effectively will not only improve language pedagogy but also release the full potential of data-driven learning. Integrating corpus linguistics into conventional teaching methods will help more efficient and evidence-based language instruction as educational systems change.

References

1. Friginal, E., Dye, P., & Nolen, M. (2025). Corpus-Based Approaches in Language Teaching: Outcomes, Observations, and Teacher Perspectives. Retrieved from Dergipark.
2. Ma, Q. (2025). A Corpus-Based Language Pedagogy for Pre-Service and In-Service Teachers: Theory, Research and Practice. Retrieved from UQ Languages & Cultures.
3. Mukherjee, J. (2004). Bridging the Gap between Applied Corpus Linguistics and the Reality of English Language Teaching. University of Giessen.
4. Pérez-Paredes, P. (2023). Language Corpus and Data-Driven Learning in Language Classrooms: A Systematic Review. Retrieved from PMC.