# CORPUS SOFTARE IN TECHING: EXAMIATION OF LANGUAGE EXPOSURE

**Olimboyeva Parizoda**

4th course student of Uzbekistan
State World Languages University

**Abstract:** In the dynamic field of English as a Foreign Language (EFL) education, corpus-based instruction (CBI) has risen as a transformative approach, enriching language learning and teaching through its diverse applications and a solid theoretical grounding. In linguistic research, corpus linguistics involves the collection and analysis of authentic texts to provide evidence for describing the nature, structure, and use of languages.

In the 1990s, European linguists and educators proposed the use of corpora as an auxiliary means for foreign language teaching. Corpus linguistics, an important branch of corpus studies, can serve as an auxiliary means for foreign language teaching and is considered an effective teaching method. Leech argues that the utilization of corpora in language teaching can occur in two distinct ways: either indirectly or directly. The indirect use of corpora refers to the application of corpus data in reference publishing, materials development, and language testing, such as the creation of dictionaries, syllabi, and teaching materials, and the construction, compilation, and selection of language tests. The direct use of corpora involves the integration of corpus data in the actual teaching process through "teaching about, teaching to exploit, and exploiting to teach". In recent years, corpora have also become a valuable resource in the field of language teaching and learning, positively impacting curriculum design, testing, and material development. It provides authentic language input, enables evidence-based teaching, and fosters learner

autonomy by offering students opportunities to scrutinize corpus data, formulate hypotheses, and develop rules to gain inductive insights into language. By using corpora, teachers can provide learners with more accurate and reliable information about language use and structure, while also promoting active learning and engagement.

In addition, there is closely relationship between Corpus-Based Instruction (CBI) and data-Driven Learning (DDL), with both methods relying on authentic language data from corpora. Data-Driven Learning (DDL) involves learners directly interacting with corpus data to discover linguistic patterns and rules. This approach encourages inductive learning and fosters greater learner autonomy and analytical skills. While CBI is typically teacher-led, using corpus data to inform lesson content and instruction, DDL places learners in the role of language researchers, engaging them directly with corpus analysis to promote discovery-based learning. CBI leverages large-scale electronic texts (corpora) to support language teaching by analyzing real language usage data. By presenting examples of actual usage, CBI encourages learners to inductively discover language rules and conventions through the analysis of real contexts, leading to more effective language acquisition. Furthermore, integrating corpus-based approaches into language teaching has the potential to revolutionize grammar instruction by providing empirical evidence of language use. CBI enables teachers to demonstrate language features as they occur in real-life contexts, bridging the gap between theoretical descriptions and practical usage, thereby enhancing learners' grammatical competence and communicative skills. The integration of CBI and DDL offers a comprehensive approach to language education. CBI provides structured exposure to authentic language use under the guidance of teachers, while DDL empowers learners to take charge of their own learning through corpus exploration and analysis. Incorporating corpus data into language teaching not only offers authentic examples of language use but also highlights linguistic variations and changes, providing learners with tools for their own linguistic investigations. Thus, CBI and DDL together support a more learner-centered and exploratory approach to language learning.

Early research on corpus-based instruction primarily focused on vocabulary and lexical competence, demonstrating its effectiveness in expanding learners' vocabulary through authentic language data. This approach enhances lexical awareness and retention by exploring word meanings and collocations. Building upon this foundation, CBI has revolutionized grammar learning and instruction by using corpora to exemplify grammatical structures with real-world examples, aiding in the understanding of language mechanics and fostering a more intuitive grasp of grammar. Additionally, CBI has been shown to substantially improve students' writing abilities, particularly in academic contexts, by providing models of proficient language use and supporting the development of syntactic awareness and well-organized texts. The integration of reading materials into writing tasks has further bolstered language exposure and writing proficiency, highlighting the mutually beneficial relationship between these skills. In the realm of collocation learning, CBI has proven advantageous in teaching conventional word pairings, essential for achieving fluency and naturalness in language use. Moreover, CBI has significantly enhanced pragmatics and communicative competence by improving learners' ability to produce pragmatic routines through explicit instruction and authentic language materials. Furthermore, CBI fosters learner autonomy and self-directed learning by empowering learners to take charge of their educational journey through engaging with corpora. Broadly, the application of corpus linguistics in language teaching is essential for developing teaching materials and enhancing discourse awareness, providing a more empirical and autonomous approach to language learning.

Despite the increasing research on the use of corpora in classrooms, our understanding of what truly happens in the classroom remains relatively lacking. Against this backdrop, the research method adopted a systematic literature review, conducting extensive searches of relevant databases and analyzing academic literature that met the criteria. The systematic review aims to thoroughly analyze the empirical studies on corpus-based instruction (CBI) in EFL classrooms, exploring research contexts, theoretical

foundations, types of classroom activities, research methods, corpus tools, influencing factors, and the advantages and challenges of implementing CBI. It also proposes effective coping strategies. To establish the research context of CBI in EFL classrooms, this systematic review first analyzes the distribution of empirical studies across three key dimensions: time, geographic regions, and affiliated research institutions. These aspects provide insights into the evolution, global reach, and institutional support of CBI research, setting the stage for a deeper exploration of its theoretical foundations, classroom practices, and outcomes.

This systematic review seeks to provide a comprehensive analysis of the application of Corpus-Based Instruction in English as a Foreign Language classrooms. The analysis, addressing a series of research questions, reveals that CBI has significant potential for enhancing students' English proficiency, including vocabulary, grammar, writing, reading, and communicative skills, as well as fostering students' autonomous learning abilities. However, challenges and limitations in CBI implementation are also evident. While various classroom activities have been developed to target different skills and purposes, most studies predominantly focus on writing and vocabulary development among proficient learners in higher education. This uneven focus on specific proficiency levels restricts the broader application of CBI. What's more, the complexity and time-consuming nature of corpus use necessitate initial training and support for both teachers and learners. This requirement may be challenging for lower-level students, who could struggle with understanding and analyzing data and may become overly reliant on corpus data. Instructional design and implementation also present challenges. In response, the review suggests several effective strategies to address these issues, including comprehensive training, simplified resources and tasks, personalized learning paths, and increased resource allocation. These strategies aim to enhance the effectiveness of CBI in EFL classrooms and improve language learning outcomes.

# REFERENCES:

1. Cobb, T. (n.d.). Compleat Lex Tutor v.8.5 [Software]. Accessed 17 July 2022 at https://www.lextutor.ca (https://www.lextutor.ca/)

2. Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), The Cambridge handbook of English corpus linguistics (pp. 478-497). Cambridge University Press. https://doi.org/10.1017/CBO9781139764377.027

3. Collentine, J. (2000). Insights into the construction of grammatical knowledge provided by user-behavior tracking technologies. Language Learning & Technology, 3(2), 44-57. https://doi.org/10125/25072

4. Corino, E., & Onesti, C. (2019). Data-Driven Learning: A Scaffolding Methodology for CLIL and LSP Teaching and Learning. Frontiers in Education, 4(7), 1-12. https://doi.org/10.3389/ (https://doi.org/10.3389) feduc.2019.00007

5. Council of Europe. (2001). Common European Framework of Reference for Languages: Learning, teaching, assessment. Cambridge University Press. Accessed 16 July 2022 at https://rm.coe.int/1680459f97

6. Crossley, S. A., Bradfield, F., & Bustamante, A. (2019). Using human judgments to examine the validity of automated grammar, syntax, and mechanical errors in writing. Journal of Writing Research, 11(2), 251- 270. https://doi.org/10.17239/jowr-2019.11.02.01

7. Crosthwaite, P. (2017). Retesting the limits of data-driven learning: feedback and error correction, Computer Assisted Language Learning, 30(6), 447-473. https://doi.org/10.1080/09588221.2017.1312462

8. Crosthwaite, P. (2020). Data-driven learning for the next generation: Corpora and DDL for pre-tertiary learners.

9. Crosthwaite, P., & Cheung, L. (2019). Learning the Language of Dentistry: Disciplinary Corpora in the Teaching of English for Specific Academic Purposes. John Benjamins. https://doi.org/10.1075/scl.93

10. Cushing, S. T. (2017). Corpus linguistics in language testing research. Language Testing, 34(4), 441-449. https://doi.org/10.1177/0265532217713044

11. De Smet, M. J. R., Leijten, M., & Van Waes, L. (2018). Exploring the process of reading during writing using eye tracking and keystroke logging. Written Communication, 35(4), 411447. https://doi.org/10.1177/0741088318788070

12. Edmonds, P. (2013). Just The Word. Accessed 17 July 2022 at http://www.just-the-word.com/

13. Flowerdew, J. (2009). Corpora in Language Teaching. In M. H. Long & C. J. Doughty (Eds.), The handbook of language teaching (pp. 327-350). Wiley-Blackwell. https://doi.org/10.1002/9781444315783.ch19

14. Flowerdew, L. (2010). Using corpora for writing instruction. In A. O'Keeffe, & M. McCarthy (Eds.). The Routledge handbook of corpus linguistics (pp. 444-457). Routledge. https://www.routledgehandbooks.com/doi/10.4324/9780203856949.ch32

15. Flowerdew, L. (2015). Data-driven learning and language learning theories: Whither the twain shall meet. In A. Leńko-Szymańska & A. Boulton (Eds.), Multiple affordances of language corpora for data-driven learning (pp. 15–36). John Benjamins. https://doi.org/10.1075/scl.69.02flo

16. Frankenberg-Garcia, A., Rees, G., Lew, R., Roberts, J., Sharma, N., & Butcher, P. (2019). ColloCaid: a tool to help academic English writers find the words they need. In F. Meunier (Eds.), CALL and complexity – short papers from EUROCALL 2019 (pp.144–150). https://doi.org/10.14705/rpnet.2019.38.1000

17. Gilquin, G., & Granger, S. (2022). 'Using data-driven learning in language teaching'. In A. O'Keeffe, & M. McCarthy (Eds.) The Routledge handbook of corpus linguistics. Second Edition (pp. 430-442). Routledge. https://doi.org/10.4324/9780367076399-30

18. Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. Behavior Research Methods, Instruments, & Computers, 36(2), 193-202. https://doi.org/10.3758/BF03195564

19. Granger, S. (1994). The Learner Corpus: A revolution in applied linguistics. English Today, 10(3), 25-33. https://doi.org/10.1017/S0266078400007665

20. Granger, S. (2002). A bird's-eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), Computer learner corpora, second language acquisition and foreign language teaching (pp. 3–33). John Benjamins. https://doi.org/10.1075/lllt.6.04gra

21. Granger, S. (2015). The contribution of learner corpora to reference and instructional materials design. The Cambridge Handbook of Learner Corpus Research, 485-510. https://doi.org/10.1017/cbo9781139649414.022

22. Granger, S., & Meunier, F. (1994). Towards a grammar checker for learners of English. In U. Fries, & G. Tottie (Eds.) Creating and using English language corpora (pp. 79-91). Rodopi.