Materiallar to'plami · Conference Proceedings
"ZAMONAVIY TADQIQOTLAR VA BARQAROR RIVOJLANISH" respublika ilmiy-amaliy konferensiyasi
04 August 2026
UDK 37
PEDAGOGIKA

TYPES OF CORPUS TECHNOLOGIES IN ESP CLASSROOM:APPLICATIONS AND PEDAGOGICAL IMPLICATIONS

Manzura Ne'matjon qizi Mamajonova*
PhD student, Uzbekistan State World Languages University
✉ manzurawebster94@gmail.com

ABSTRACT

Corpus technologies have become an essential component of English for Specific Purposes (ESP) instruction by providing learners with access to authentic language used in professional and academic contexts. This article explores the major types of corpus technologies employed in ESP classrooms mainly for IT sphere, including general reference corpora, specialized corpora, learner corpora, web-based corpora, concordancing software, and corpus query systems. It examines their pedagogical functions and demonstrates how these technologies facilitate the development of learners' lexical, grammatical, discourse, and communicative competence through data-driven learning (DDL). Drawing on recent studies in corpus linguistics and computer-assisted language learning, the article discusses the advantages and limitations of different corpus technologies and their application in designing learner-centered ESP activities. The article concludes that corpus technologies should be considered an integral element of modern ESP pedagogy and recommends expanding teacher training and corpus-based instructional materials to maximize their educational potential.

Keywords

corpus technologies, corpus linguistics, English for Specific Purposes (ESP), computer-assisted language learning (CALL), data-driven learning, concordancers, specialized corpora, language teaching.

DOI: 10.5281/zenodo.21743742

INTRODUCTION

English for Specific Purposes (ESP) has become one of the most dynamic branches of language education, responding to the increasing demand for English proficiency in specialized academic and professional domains such as engineering, information technology, medicine, business, and law. Unlike General English, ESP focuses on developing learners' communicative competence within discipline-specific contexts, enabling them to perform professional tasks effectively through appropriate language use. Consequently, contemporary ESP pedagogy emphasizes authentic language input, learner autonomy, and context-based instruction rather than the isolated teaching of grammar and vocabulary (Hyland, 2022; Jablonkai & Csomay, 2022).

One of the most significant developments supporting these pedagogical principles is the emergence of corpus linguistics. Corpus linguistics is the empirical study of language through electronically stored collections of authentic spoken and written texts known as corpora. Since the pioneering work of Sinclair (1991), corpus linguistics has transformed language teaching by allowing researchers and educators to investigate how language is naturally used across different communicative contexts. Instead of relying solely on intuition or textbook examples, corpus technologies provide authentic linguistic evidence that enables learners to observe recurring lexical, grammatical, pragmatic, and discourse patterns in real-life communication (McEnery & Hardie, 2012).

Recent advances in digital technologies have considerably expanded the range of corpus resources available for educational purposes. Modern ESP classrooms can benefit from various corpus technologies, including general reference corpora such as the Corpus of Contemporary American English (COCA) and the British Na-

tional Corpus (BNC), specialized disciplinary corpora, learner corpora, web-based corpora, concordancing software such as AntConc and LancsBox, and comprehensive corpus management platforms including Sketch Engine. These technologies provide learners with opportunities to explore authentic language independently, compare linguistic patterns across genres, and discover discipline-specific expressions through data-driven learning (DDL) (Boulton & Cobb, 2017; Pérez-Paredes, 2022).

Table 1. Classification of Corpus Technologies Used in ESP Instruction

Corpus Technology	Representative Resources	Pedagogical Applications in ESP	Main Advantages
General Reference Corpora	COCA, BNC, iWeb	Teaching vocabulary, collocations, grammar, register differences, academic and professional writing	Large datasets, authentic examples, register comparison
Specialized ESP Corpora	BAWE, MICASE, Engineering Corpus, Business Letter Corpus, Medical Research Corpus	Teaching discipline-specific terminology, genre conventions, academic writing, professional communication	Authentic domain-specific language, increased relevance for ESP learners
Learner Corpora	ICLE, LOCNESS, EFCAMDAT	Error analysis, contrastive analysis, feedback, materials development	Identifies frequent learner difficulties and supports targeted instruction
Parallel Corpora	Europarl, OpenSubtitles, OPUS	Translation studies, bilingual ESP instruction, terminology comparison	Facilitates cross-linguistic awareness and translation competence
Web-based Corpora	GloWbE, NOW Corpus	Teaching contemporary vocabulary, professional discourse, global English varieties	Large-scale, up-to-date, multilingual language data
Concordancing Software	AntConc, LancsBox	Data-driven learning (DDL), collocation analysis, phraseology, discourse analysis	Free or low-cost, promotes learner autonomy and inductive learning
Corpus Query Systems	Sketch Engine	Word sketches, keyword analysis, collocations, lexical bundles, semantic relations	Powerful analytical tools for teachers and researchers
AI-enhanced Corpus Technologies	Sketch Engine AI features, intelligent corpus interfaces	Automated feedback, corpus-assisted writing, personalized language learning	Faster analysis, adaptive learning support, enhanced accessibility

Data-driven learning has become one of the most influential corpus-based teaching approaches in second language acquisition. Rather than receiving linguistic rules directly from teachers, learners investigate concordance lines, frequency lists, collocations, and lexical bundles to formulate their own conclusions regarding language use. This inductive learning process promotes deeper cognitive engagement, enhances noticing of linguistic features, and encourages learner autonomy. Numerous studies have demonstrated that corpus-assisted learning contributes positively to vocabulary acquisition, academic writing, phraseological competence, grammatical accuracy, and disciplinary literacy in ESP contexts (Chen & Flowerdew, 2018; Vyatkina, 2020).

Despite the growing popularity of corpus-based instruction, many ESP teachers remain uncertain about selecting appropriate corpus technologies for different instructional objectives. While general corpora are useful for investigating high-frequency language and register variation, specialized corpora provide discipline-specific terminology and genre conventions that more accurately reflect learners' future professional communication. Concordancing programs facilitate detailed linguistic analysis, whereas corpus query systems offer sophisticated functions such as collocation analysis, word sketches, keyword extraction, and semantic profiling. Furthermore, recent developments integrating artificial intelligence with corpus technologies have enhanced corpus accessibility by enabling automatic annotation, intelligent search functions, and personalized learning environments (Pérez-Paredes, 2022).

Most investigations focus on individual corpus tools or specific language skills rather than comparing multiple corpus technologies within a unified instructional framework. As a result, ESP practitioners often lack comprehensive guidance regarding which corpus resources are most appropriate for particular learning objectives and classroom contexts.

Appropriate chosen corpus tools not only provide teachers with necessary method but also bring into classroom nativeness (Radjabova, 2025). Moreover, corpus-based tasks create opportunity to practice authentic IT discourse which can be effectively used by instructors (Nesi, H., & Gardner, S. 2012).

Therefore, this article aims to classify and analyze the principal types of corpus technologies used in contemporary ESP classrooms and to examine their pedagogical applications in language teaching. It reviews current literature on corpus linguistics and data-driven learning, compares widely used corpus resources according to their educational functions, and discusses the advantages and limitations of each technology. By synthesizing recent developments in corpus-assisted language teaching, the article seeks to provide practical recommendations for ESP educators interested in integrating corpus technologies into learner-centered instruction and promoting authentic, evidence-based language learning.

CONCLUSION

In conclusion, corpus technologies play a significant role in modern ESP instruction by providing authentic, discipline-specific language resources that enhance learners' linguistic and communicative competence. The classification presented in this article demonstrates that different corpus technologies including corpus resources, corpus analysis tools, and AI-enhanced corpus platforms serve distinct yet complementary functions in supporting vocabulary development, genre awareness, discourse analysis, and data-driven learning. Their effective integration into ESP classrooms encourages learner autonomy, critical language awareness, and evidence-based language learning while enabling teachers to design more authentic and professionally relevant instructional activities. As corpus technologies continue to evolve alongside artificial intelligence, they are expected to become increasingly accessible and influential in preparing ESP learners for successful communication in academic and professional environments.

REFERENCES

1. Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, *67*(2), 348–393. https://doi.org/10.1111/lang.12224
2. Chen, M., & Flowerdew, J. (2018). A critical review of research and practice in data-driven learning (DDL) in the academic writing classroom. *International Journal of Corpus Linguistics*, *23*(3), 335–369. https://doi.org/10.1075/ijcl.16127.che
3. Hyland, K. (2022). English for specific purposes: What is it and where is it taking us? *ESP Today*, *10*(2), 202–220.
4. Jablonkai, R. R., & Csomay, E. (Eds.). (2022). *The Routledge handbook of corpora and English language teaching and learning.* Routledge.
5. McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice.* Cambridge University Press.
6. Nesi, H., & Gardner, S. (2012). Genres across the disciplines: Student writing in higher education. Cambridge University Press.
7. Pérez-Paredes, P. (2022). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2020. *Computer Assisted Language Learning*, *35*(9), 2295–2325. https://doi.org/10.1080/09588221.2021.1892768
8. Radjabova, G. (2025). Using corpus of contemporary american english [coca] to enhance writing competence of b2 level students. *Foreign Languages in Uzbekistan*, 110-124.
9. Sinclair, J. (1991). *Corpus, concordance, collocation.* Oxford University Press.
10. Vyatkina, N. (2020). Data-driven learning for language pedagogy: Toward the development of second language corpus literacy. *Language Learning & Technology*, *24*(3), 1–15.