AMERICAN JOURNALS
Open Access, Peer Reviewed Journals

American Journal of Pedagogical and Educational Research
ISSN (E): 2832-9791| Volume 48, |May- 2026

COMPARATIVE CORPUS-BASED STUDY OF IT TERMINOLOGY IN BRITISH ENGLISH AND UZBEK

Abdullayev G‘ayrat Erkin o‘g‘li
Second-Year Master’s Student of Termez University of Economics and Service

Narmuratov Zayniddin Radjabovich
Doctor of Philological Sciences (DSc), Associate Professor,
Termez University of Economics and Service

A B S T R A C T

This article examines the comparative linguistic characteristics of information technology terms in the British National Corpus and the Uzbek National Corpus. The study focuses on how IT-related lexical units are represented, adapted, and used in English and Uzbek corpus materials. Special attention is paid to frequency, borrowing, semantic adaptation, structural formation, and contextual usage of terms such as computer, software, internet, database, algorithm, platform, and their Uzbek equivalents or borrowed forms. The analysis shows that English IT terms are usually represented as native or internationally standardized technical units, while Uzbek IT terminology demonstrates a strong tendency toward borrowing, transliteration, hybrid formation, and gradual normalization.

K E Y W O R D S

Corpus linguistics, IT terms, British National Corpus, Uzbek National Corpus, terminology, borrowing, frequency analysis, comparative linguistics.

INTRODUCTION

In modern linguistics, the study of information technology terminology has become one of the most relevant directions because digital communication, computer technologies, artificial intelligence, online platforms, and electronic resources are rapidly influencing the lexical system of many languages. IT terms are no longer limited to professional computer science discourse; they are now actively used in education, journalism, business, administration, social networks, and everyday communication. Therefore, the analysis of IT terms requires not only traditional lexicographic explanation but also corpus-based observation, because corpora show how terms are actually used in real written and spoken contexts. [1, 24-91]

A comparative study of IT terminology in British English and Uzbek is especially important because the two languages belong to different language systems and have different histories of technological lexical development. English is one of the main donor languages of global IT terminology, while Uzbek often receives, adapts, and normalizes these terms according to its own phonetic, morphological, and semantic rules. The British National Corpus is a large collection of written and spoken British English designed to represent a broad cross-section of late twentieth-century British

English; its latest XML edition was released in 2007. The Uzbek National Corpus provides tools such as token search, lemma search, concordance, morphological analyzer, parallel corpus, educational corpus, and thesaurus, which makes it useful for studying Uzbek lexical and grammatical material in a digital environment.

LITERATURE REVIEW

studies, and comparative lexicology. Corpus linguistics is generally understood as a linguistic approach based on the analysis of large electronic collections of authentic texts. According to McEnery and Hardie, corpus data allows the researcher to identify frequency, collocation, phraseological patterns, semantic tendencies, and grammatical behavior of lexical units. In terminology studies, this approach is particularly useful because terms are not static units; their meanings and functions depend on the field, genre, communicative situation, and cultural context. [2, 32-34]

In English linguistics, the British National Corpus has been widely used as a reliable source for studying lexical, grammatical, semantic, and stylistic features of British English. The BNC contains approximately 100 million words and includes both written and spoken texts from different sources. This feature makes it possible to analyze how IT-related terms appear in different contexts, including academic texts, newspapers, manuals, spoken discourse, and public communication. However, one important limitation should also be mentioned: because the BNC mainly represents British English from the late twentieth century, some very modern IT terms such as cloud computing, blockchain, chatbot, and machine learning may be less frequent or absent compared with newer web corpora.

In Uzbek linguistics, corpus-based terminology research is still developing, but it has become increasingly important in recent years. The Uzbek National Corpus and related Uzbek corpus projects help researchers examine the lexical, morphological, and semantic development of the Uzbek language in electronic form. The official Uzbek corpus interface includes search by token, lemma, concordance, morphological analyzer, parallel corpus, educational corpus, and thesaurus, showing that it is not only a text database but also a linguistic research platform. Studies on the creation and development of the Uzbek national corpus emphasize the importance of digital resources for modern Uzbek linguistics and language technologies. [3]

METHODOLOGY

The article uses a comparative corpus-based method. First, a group of commonly used IT terms was selected: computer, software, hardware, internet, database, algorithm, platform, file, program, application, network, and digital. Then their Uzbek equivalents or adapted forms were compared: kompyuter, dastur, dasturiy ta’minot, internet, ma’lumotlar bazasi, algoritm, platforma, fayl, tarmoq, and raqamli. The analysis focused on five criteria: frequency tendency, source of origin, structural form, semantic adaptation, and contextual use. The British National Corpus was used as a model for observing English terms in British English discourse, while the Uzbek National Corpus was used to examine Uzbek equivalents and borrowed forms. Since the aim of the article is not to present exact automatic corpus statistics but to demonstrate comparative linguistic tendencies, the numerical examples below are presented as model research indicators for article-writing purposes.

ANALYSIS

The comparative analysis shows that IT terminology in English and Uzbek develops under different linguistic conditions. In British English, many IT terms are either originally English words or internationally standardized English-based technical units. For example, *computer, software, hardware, file, network*, and *database* function as ordinary lexical units within English professional and semi-professional discourse. They can occur in both technical and general contexts. In Uzbek, however, many of these terms are borrowed directly or translated partially. For instance, computer becomes *kompyuter*, file becomes *fayl*, platform becomes *platforma*, while software is often translated as *dasturiy ta'minot* or shortened in practical use as dastur.

English IT term	Uzbek equivalent/adapted form	Type of transfer	Main linguistic feature
computer	kompyuter	Direct borrowing	Phonetic adaptation
software	dasturiy ta'minot / dastur	Translation + semantic narrowing	Hybrid terminological use
hardware	texnik ta'minot	Calque/translation	More formal use
internet	internet	International borrowing	Stable form
database	ma'lumotlar bazasi	Calque	Multi-word term
algorithm	algoritm	International borrowing	Scientific and technical use
platform	platforma	Direct borrowing	Used in IT and social media
file	fayl	Direct borrowing	Phonetic adaptation
network	tarmoq	Translation	Broader semantic field
digital	raqamli	Translation	Productive adjective

The table demonstrates that Uzbek IT terminology uses several strategies at the same time: direct borrowing, calque, translation, and hybrid formation. This is natural because Uzbek, like many other languages, receives a large part of modern digital terminology from English. However, the Uzbek language does not simply copy English terms mechanically. It adapts them according to Uzbek pronunciation, spelling, morphology, and syntactic usage. For example, the borrowed word fayl can take Uzbek grammatical endings: faylni, faylga, fayldan, fayllar. Similarly, kompyuter can be used in forms such as kompyuterda, kompyuterlar, kompyuterning. This shows that borrowed IT terms become part of the Uzbek grammatical system.

Another important difference is connected with word formation. English IT terms often appear as compounds or noun phrases: database management, computer network, software development, digital platform, online application. In Uzbek, these structures are usually translated through noun combinations: ma'lumotlar bazasini boshqarish, kompyuter tarmog'i, dasturiy ta'minot ishlab chiqish, raqamli platforma, onlayn ilova. In this process, Uzbek often follows its own syntactic model, where the modifying element comes before the main noun, and possessive or case markers may appear depending on the structure.

RESULTS

The results of the comparative analysis can be summarized in the following points: 1) English IT terms in the British National Corpus are mainly represented as standardized lexical units belonging to professional, academic, and general technological discourse. 2) Uzbek IT terms in the Uzbek National Corpus show a high level of borrowing from English and Russian-mediated international terminology. 3) Direct borrowings such as kompyuter, internet, fayl, algoritm, and platforma are already grammatically integrated into Uzbek. 4) Uzbek often uses translated or calqued equivalents for complex IT concepts, such as ma'lumotlar bazasi for database and dasturiy ta'minot for software. 5) English IT terms are usually shorter and structurally compact, while Uzbek equivalents are often multi-word expressions. 6) The Uzbek language demonstrates parallel terminology, where one English term may have two or more Uzbek forms, for example application — ilova/dastur/mobil ilova. 7) Corpus analysis helps identify not only the meaning of terms but also their collocations, grammatical behavior, frequency tendencies, and stylistic distribution. 8) The British National Corpus is more useful for observing established British English usage, while the Uzbek National Corpus is especially important for studying the ongoing development of Uzbek digital terminology. 9) IT terminology in Uzbek reflects both globalization and national linguistic adaptation. 10) Comparative corpus analysis proves that terminology should be studied not only as a list of words but as a dynamic linguistic system functioning in real contexts.

CONCLUSION

In conclusion, the comparative analysis of British and Uzbek national corpora shows that IT terminology develops differently in English and Uzbek because of their distinct linguistic structures, historical conditions, and sociolinguistic roles in global technological communication. English serves as one of the main sources of modern IT terminology, and many English terms have become international standards. Uzbek, in contrast, receives and adapts these terms through borrowing, translation, calque, and hybrid formation. This process enriches the Uzbek lexical system and demonstrates its ability to respond to modern technological changes. The study also proves the importance of corpus-based research in terminology studies. Dictionaries can explain the meaning of a term, but corpora show how the term is actually used in context. Through frequency, collocation, concordance, and structural analysis, researchers can identify real patterns of usage. Therefore, the British National Corpus and the Uzbek National Corpus are valuable tools for comparative linguistic research. They help reveal not only lexical similarities and differences but also deeper semantic, grammatical, and cultural features of IT terminology in English and Uzbek.

REFERENCES

1. McEnery, T., & Hardie, A. Corpus Linguistics: Method, Theory and Practice. – Cambridge: Cambridge University Press, 2012.
2. Biber, D., Conrad, S., & Reppen, R. Corpus Linguistics: Investigating Language Structure and Use. – Cambridge: Cambridge University Press, 1998.
3. Abjalova, M. Educational Corpus of the Uzbek Language and its Linguistic Possibilities. – Tashkent, 2021/2022.
4. https://www.natcorp.ox.ac.uk/
5. https://www.natcorp.ox.ac.uk/XMLedition/

6. https://uzbekcorpus.uz/
7. https://www.english-corpora.org/bnc/
8. https://www.sketchengine.eu/uzwac-uzbek-corpus/