

## EFFICIENCY OF USING CHATGPT MODELS IN LINGUISTIC ANALYSIS

**Ruxshona Rajabova**

Student, Group 6-4 Uzb 25  
Bukhara State University Uzbekistan

**Abstract:** This article examines the effectiveness of applying artificial intelligence technologies in linguistic analysis, with particular emphasis on the use of ChatGPT models. The study explores innovative approaches to morphological, syntactic, semantic, lexical, and corpus-based text analysis using large language models and compares these methods with traditional linguistic analysis techniques. Special attention is given to the model's ability to understand the contextual and cultural features of the Uzbek language and to its applicability in scientific linguistic research. The findings demonstrate that ChatGPT models significantly enhance the speed, accuracy, and interactivity of linguistic analysis. At the same time, the study identifies the need for further adaptation of large language models to the structural and cultural characteristics of the Uzbek language. Overall, the article highlights new perspectives and research directions for the integration of artificial intelligence tools into modern linguistics.

**Keywords.** artificial intelligence, linguistics, ChatGPT, corpus analysis, morphological analysis, syntactic analysis, semantic analysis, Uzbek language, large language models

**Аннотация.** В данной статье подробно раскрыты возможности применения технологий искусственного интеллекта в лингвистике, в частности, различные инновационные методы, основанные на использовании модели ChatGPT. Изучается процесс морфологического, синтаксического и семантического анализа текстов с помощью модели ChatGPT, а также проводится сравнение с традиционными методами анализа. Кроме того, рассматриваются способности модели к пониманию контекста узбекского языка и возможности применения результатов анализа в научных исследованиях. Использование моделей ChatGPT позволяет проводить лингвистические анализы быстро, точно и интерактивно. Вместе с тем выявлена необходимость дальнейшей адаптации модели к особенностям узбекского языка. Также статья способствует раскрытию новых направлений использования инструментов искусственного интеллекта в лингвистике.

**Ключевые слова.** искусственный интеллект, лингвистика, ChatGPT, морфологический анализ, синтаксический анализ, семантический анализ, узбекский язык, понимание контекста.

**Introduction.** Modern linguistics is developing in close connection with digital technologies. Digital linguistics, that is, the study of language through technological and computational tools, has become a highly significant field in recent years. This area is particularly relevant for Uzbek linguistics, as artificial intelligence-based approaches offer substantial opportunities for the development of the Uzbek language, its integration into multilingual systems, the creation of linguistic corpora, and the development of automatic translation tools, especially through bilingual methodologies [3].

Alan Turing is considered one of the pioneers of artificial intelligence research. Artificial intelligence was formally established as an independent scientific field in 1956. In the summer of that year, at a conference held at Dartmouth College, John McCarthy first introduced the term "artificial intelligence," thereby entering history as the author of this concept. Although research in artificial intelligence has been conducted since the mid-twentieth century, public interest in the

field increased significantly after 2012, when deep learning demonstrated superiority over other artificial intelligence methods, and further intensified following the major achievements in transformer architectures in 2017. In the early 2020s, the field experienced rapid growth, with numerous companies, universities, and research laboratories achieving notable progress in artificial intelligence research [1].

Traditionally, linguistics has focused on analyzing language structure and its various levels, including vocabulary, phonetic features, syntactic organization, semantic context, and pragmatic aspects. Previously, such analyses were conducted manually or with the help of simple computer programs, resulting in relatively slow research processes. By the late twentieth and early twenty-first centuries, corpus linguistics emerged, followed by the introduction of stochastic models, machine learning techniques, and artificial intelligence, which significantly expanded the scope and methodological capabilities of linguistic research. Artificial intelligence is a branch of computer science dedicated to creating systems capable of performing tasks that require human intelligence. In recent decades, AI technologies have been actively integrated into educational institutions, leading to substantial changes in teaching and learning approaches. Modern technologies not only automate routine processes but also enable the personalization of educational programs according to learners' individual needs. Research indicates that the application of artificial intelligence in education can improve both the efficiency and quality of learning. For instance, AI-based platforms that utilize machine learning algorithms are able to analyze learners' knowledge levels and provide tailored tasks by taking into account their strengths and weaknesses. Additionally, interactive tools such as quizzes, flashcards, and educational games facilitate concise and accessible explanations, supporting effective knowledge consolidation.

### **Semantic Analysis and Cultural Specificity**

Language is not merely a collection of words; it also reflects the history, traditions, religious beliefs, emotions, and psychological characteristics of an entire nation. Although artificial intelligence (AI) algorithms perform accurately in tasks such as identifying lexical relationships, frequency analysis, and translation based on parallel corpora, processes related to the interpretation of culturally specific terms, ironic expressions, and implicit meanings (subtext) remain consistently problematic. For instance, metaphorical expressions commonly used in everyday speech, such as idioms and culturally bound phrases, may be translated literally by artificial intelligence systems if they are not explicitly explained, leading to semantic distortion. In such cases, precise linguistic commentary and interpretation remain essential. Undoubtedly, one of the major challenges in the application of AI in linguistics is the shortage of specialists who possess both digital competencies in linguistics and, conversely, programmers in the IT field who have sufficient linguistic knowledge. In the future, the demand for integrated specialists—professionals with solid expertise in computational linguistics, strong programming skills, and a thorough understanding of theoretical linguistics—is expected to increase. Such specialists will not only be responsible for training and improving AI models but will also contribute significantly to the development of new directions within linguistic research [2].

In the future, linguistics research driven by artificial intelligence may open new directions and opportunities. In particular, the development of universal translation systems capable of covering a wide range of languages has already begun in technologically advanced countries. Artificial intelligence-based language learning tools tailored for bilingual individuals are expected to be developed. Interactive models designed to accelerate the language learning process will emerge,

and multilingual artificial intelligence systems will enable more natural and context-aware communication frameworks [6].

### **Corpus Analysis**

In linguistics, corpus analysis refers to the systematic study of a language's structure, vocabulary, syntax, and semantic features through large-scale collections of texts known as corpora. Corpora serve as primary data sources in linguistic research, enabling the identification of word and phrase frequency, their interrelationships, sentence structure, and stylistic characteristics.

The main objectives of corpus analysis include the following:

Word Frequency Analysis — the identification of how often words and phrases occur within a text.

1. Collocation Analysis — the examination of the frequency with which words co-occur, revealing their semantic associations.
2. Syntactic and Grammatical Analysis — the analysis of sentence structure, including relationships between verbs and nouns, as well as the identification of conjunctions and other grammatical elements.
3. Stylistic and Genre Analysis — the identification of a text's style and genre, such as distinguishing between scientific and literary texts.

Large language models, particularly models such as ChatGPT, significantly enhance the efficiency of corpus analysis. These models make it possible to analyze millions of words and sentences within a short period of time, identify contextual relationships between words, and construct semantic networks. In addition, large language models (LLMs) enable the automatic annotation of texts, that is, the classification of words and sentences into linguistic categories. For example, the automatic identification of nouns, verbs, adjectives, and metaphorical expressions can substantially accelerate the process of linguistic analysis. Corpus analysis is therefore widely applied in practical linguistic research. For example, in literary analysis, corpora of nineteenth-century poetry are used to identify lexical richness and metaphorical expressions, while studies of language evolution rely on historical texts to observe changes in word meanings over time. In addition, corpus analysis enables the identification of words and expressions used across different regions and dialects. As an essential tool for linguistic research, corpus analysis significantly benefits from the use of large language models (LLMs), which greatly enhance its efficiency and accuracy. This approach not only allows for rapid text analysis but also provides deeper insights into the semantic, syntactic, and stylistic features of language [7].

### **Lexical Analysis**

Lexical analysis is a type of linguistic analysis focused on the study of words and their usage characteristics. Through corpus-based methods, it is possible to determine word frequency, contextual meanings, and semantic relationships between lexical units. Lexical analysis enables the examination of a language's lexical richness, the use of words in different meanings, and semantic relations such as synonymy and antonymy. For example, identifying word frequencies in a text and determining the most frequently used words are central aspects of lexical analysis [7],

### **Stylistic and Genre Analysis**

It is a linguistic analysis aimed at studying the stylistic and genre characteristics of a text. This analysis makes it possible to determine whether a text is written in a scientific, literary, formal, or informal style, as well as to examine the differences between various genres (article, essay, story, poem, etc.). The style of a text is identified through the analysis of word choice, sentence structure, and the use of metaphorical and stylistic devices. For example, scientific texts often

employ technical terms and formal constructions, whereas literary texts frequently feature metaphors and figurative expressions.

#### **Collocation Analysis**

In linguistics, collocation analysis is a type of study aimed at examining the frequency of co-occurrence and natural combinations of words. Collocations play an important role in understanding the semantic and syntactic features of a language, as they demonstrate how words naturally and in a specific pattern combine depending on the context. Using corpus analysis, the most frequently co-occurring word combinations can be identified. For example, the word “fast” is often combined with words such as “respond,” “learn,” and “walk.”

#### **Syntactic and Grammatical Analysis**

Syntactic and grammatical analysis is a type of linguistic study aimed at examining the function of words in a text based on their order, sentence structure, and grammatical rules. This analysis makes it possible to identify the structure of a language, the formation of sentences, and the syntactic relationships between words. The function of words within a sentence is determined, including the subject, complement, object, verb, and other syntactic roles. For example, in the sentence “Men kitob o‘qiyman” (“I read a book”), “Men” functions as the subject, “kitob” as the object, and “o‘qiyman” as the verb.

In conclusion, modern linguistics is developing in close connection with digital technologies, artificial intelligence, and large language models (LLMs). Digital linguistics enables the study of language with the aid of technological tools

aimed at studying and provides significant opportunities for the development of the Uzbek language, its integration into multilingual systems, the creation of corpora, and the development of automated translation tools. Corpus analysis enables the systematic study of the lexical, syntactic, semantic, and stylistic features of a language. Its main directions include: word frequency analysis, collocation analysis, syntactic and grammatical analysis, and stylistic and genre analysis. While lexical analysis serves to identify word usage and semantic relationships, collocation analysis examines natural word combinations. Syntactic and grammatical analysis helps determine sentence structure and the function of words, whereas stylistic and genre analysis defines the style and genre of a text. Artificial intelligence and large language models significantly accelerate corpus analysis, enabling automatic annotation and the identification of contextual relationships. At the same time, the expertise of human specialists remains crucial for accounting for linguistic subtleties such as cultural specificity, metaphorical expressions, and subtext. In the future, integrated specialists—professionals with advanced knowledge in computational linguistics—will play a key role in the development of both linguistics and artificial intelligence.

#### **References:**

1. **Akbarova, R.** *Theoretical aspects of the studied language: stylistics*. Namangan davlat universitetining raqamli kutubxonasi, 2024.
2. **“KOMPYUTER LINGVISTIKASI: MUAMMOLAR, YECHIM, ISTIQBOLLAR”** Xalqaro ilmiy-amaliy konferensiya, 2025.
3. **Mengliyev, B., & Hamroyeva, Sh.** *Sun'iy intellekt: tilshunoslar endi nima qiladi? // RAQAMLI TRANSFORMATSIYA SHAROITIDA SUN'IY INTELLEKTNI RIVOJLANTIRISH VA ZAMONAVIY MATEMATIKANI O'QITISH MUAMMOLARI VA YECHIMLARI*, 445. “Ma’rifat” gazetasi, 2024-yil 18-dekabr.



4. **The Essential Turing: The ideas that gave birth to the computer age (en).** Oxford, England: Clarendon Press, 2004.
5. **Tillayeva, O. B.** *The analysis of verbal collocations in English and Uzbek languages. Ta'lim innovatsiyasi va integratsiyasi*, 2025.
6. **Toshkent davlat o'zbek tili va adabiyoti universiteti ilmiy jurnali.** (2021). *Sun'iy intellekt va lingvistik model tizimlari*.
7. **Jurafsky, D., & Martin, J. H.** *Speech and Language Processing* (3rd ed. draft), 2025.