

SYNTACTIC COMPLEXITY IN AI-GENERATED AND HUMAN-WRITTEN ENGLISH TEXTS: A COMPARATIVE STUDY

Tojiboyeva Mushtariybonu Tojiboy qizi

Teacher at Uzbekistan World Languages University

<https://doi.org/10.5281/zenodo.17519947>

Annotation: This paper explores the syntactic complexity of English texts generated by Artificial Intelligence (AI) compared to those written by humans. The study aims to identify how AI structures sentences, uses subordinate clauses, and manages syntactic variation in contrast to human writing. A corpus of 40 texts—20 written by humans and 20 generated by Chat GPT—was analyzed using measures such as mean sentence length, subordination ratio, and clause density. The results reveal that AI-generated texts tend to exhibit syntactic balance and formal regularity, while human-written texts demonstrate greater syntactic variation and stylistic flexibility.

Keywords: syntactic complexity, AI-generated text, sentence structure, human writing, clause density, stylistic variation.

The rapid development of AI language models has significantly impacted written communication, raising questions about linguistic authenticity and stylistic individuality. While AI can produce grammatically correct and semantically coherent texts, there remains a noticeable difference in syntactic variation between AI-generated and human-written texts.

Syntactic complexity refers to the degree of variation and elaboration in sentence structure, which reflects the writer's linguistic competence and stylistic intent. This study investigates how AI handles complex syntax compared to human writers, focusing on clause relationships, subordination, and sentence length.

A comparative corpus consisting of 40 English texts was compiled. Twenty texts were produced by Chat GPT (GPT-5 model), and twenty were written by university-level human authors.

The analysis employed several key syntactic measures:

Mean Sentence Length (MSL) – the average number of words per sentence.

Clause Density (CD) – the ratio of clauses to sentences.

Subordination Index (SI) – the proportion of dependent clauses to total clauses.

The texts were analyzed using linguistic software tools such as L2 Syntactic Complexity Analyzer.

The results show that AI-generated texts tend to produce sentences with moderate length and consistent syntactic structure. The mean sentence length for AI-generated texts was 18.4 words, while for human texts it reached 22.7 words.

Clause density was slightly higher in human-written texts (1.53 clauses per sentence) compared to AI-generated ones (1.32).

The subordination index indicated that humans used more embedded and relative clauses, whereas AI texts showed a preference for coordination.

This difference highlights that AI systems favor clarity and syntactic predictability, whereas human writers use syntactic variation to express subtle meanings, emphasize ideas, or create rhythm.

For instance, in AI texts, sentences like “The system analyzes the data and produces a report” dominate, while human texts often contain structures such as “The system, which had previously failed to respond, now performs the analysis and produces a detailed report.”

A closer look at sentence types revealed that AI-generated texts favor balanced and predictable structures. Sentences such as “AI improves efficiency and saves time” are grammatically correct but stylistically plain. In contrast, human writers often employ embedded or participial clauses: “AI, which has dramatically reshaped the workplace, continues to improve efficiency and save valuable time.”

Human writing shows more instances of stylistic rhythm — variation in sentence length and type — contributing to textual naturalness and reader engagement.

Moreover, AI tends to maintain a formal register regardless of genre, avoiding contractions or colloquial syntax. Human authors, on the other hand, adapt syntactic choices according to communicative intent and audience expectations. The comparative analysis demonstrates that AI-generated texts exhibit less syntactic variation than human-written texts. While AI maintains grammatical accuracy and coherence, it tends to avoid deep subordination and stylistic experimentation. AI-generated texts exhibit syntactic coherence and grammatical precision, yet they fall short in mirroring the syntactic depth and stylistic variation of human writing. Humans use subordination, embedding, and stylistic rhythm to convey emphasis and nuance — strategies not yet fully mastered by AI.

In conclusion, syntactic complexity remains a strong indicator of human authorship. These findings suggest that future AI development should aim to enhance syntactic adaptability and discourse-level awareness to approximate human linguistic creativity more closely.

Adabiyotlar, References, Литературы:

1. Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5–35.
2. Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474–496.
3. OpenAI. (2024). GPT-5 Model Card. <https://openai.com/research>
4. Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis. *Studies in Second Language Acquisition*, 25(4), 497–528.