Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 04, July, 2025
Website: usajournals.org

THE IMPACT OF CHATGPT ON MODERN LINGUISTICS AND THE CHALLENGES IT POSES

Normamatova Zakhro Dilshodovna
English Teacher; Karshi, Kashkadarya;
Tel: +998979661997;
e-mail: zakhronormamatova97@icloud.com

Abstract
This article explores the influence of ChatGPT, a leading generative language model developed by OpenAI, on modern linguistics. It examines how ChatGPT is reshaping linguistic theories in syntax, semantics, pragmatics, and discourse, as well as its applications in computational linguistics and language pedagogy. At the same time, the paper discusses ethical and methodological challenges such as bias, grounding, and over-reliance. Drawing on recent scholarly perspectives, this paper aims to evaluate both the potential and limitations of large language models (LLMs) like ChatGPT in contemporary linguistic research and practice.

Keywords: ChatGPT, modern linguistics, syntax, semantics, pragmatics, computational linguistics, AI bias, large language models, language pedagogy

Introduction

The emergence of advanced language models such as ChatGPT has initiated a profound shift in how language is processed, produced, and studied. Traditionally, linguistics has examined natural language based on human cognition, usage, and structure. With the development of transformer-based models (Vaswani et al., 2017), which simulate human-like language through massive data processing, questions arise about what constitutes “language understanding” and how AI reshapes the epistemology of linguistics.

ChatGPT’s influence spans from practical applications in education and translation to theoretical debates about the nature of syntax, semantics, and

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 04, July, 2025
Website: usajournals.org

pragmatics. This paper examines these developments and investigates the opportunities and challenges posed by integrating AI into linguistic theory and practice.

2. Theoretical Shifts in Linguistic Paradigms

2.1 Syntax and Generative Grammar

ChatGPT generates syntactically fluent sentences without relying on explicit grammar rules. This challenges Noam Chomsky's theory of Universal Grammar (UG), which posits an innate human faculty for syntax (Chomsky, 1965). The model's success suggests that exposure to massive data can substitute for inborn syntactic knowledge—at least functionally. Yet, while ChatGPT produces well-formed sentences, it may lack deep grammatical competence, highlighting the distinction between performance and true linguistic knowledge (Berwick & Chomsky, 2016).

2.2 Semantics and the Illusion of Understanding

Semantics involves meaning—something ChatGPT approximates statistically. Bender and Koller (2020) argue that models like ChatGPT merely manipulate symbols without understanding them, a criticism known as the "stochastic parrot" argument. For instance, ChatGPT may correctly use a word in context but lack any real-world referent or experience to ground that usage. This poses questions for truth-conditional semantics and reinforces the gap between form and meaning.

2.3 Pragmatics and Contextual Inference

ChatGPT handles turn-taking and politeness but struggles with deeper pragmatic functions like implicature, presupposition, or irony (Searle, 1980). Its inability to hold intentions or beliefs limits its pragmatic competence. For example, while it can simulate apologies or requests, it does not truly "intend" them. Thus, while superficially pragmatic, ChatGPT lacks the cognitive foundation for full communicative competence.

3. Applications in Linguistics and Language Learning

3.1 Computational Linguistics and NLP

ChatGPT is a landmark in Natural Language Processing (NLP), outperforming earlier models in translation, summarization, and dialogue systems. It uses

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 04, July, 2025
Website: usajournals.org

transformer architectures (Vaswani et al., 2017) to model long-range dependencies in text. In computational linguistics, it has shifted focus from rule-based systems to probabilistic modeling. However, researchers warn of reduced model transparency and increased difficulty in interpreting linguistic patterns (Rogers et al., 2020).

3.2 Corpus Linguistics and Simulation

ChatGPT can generate extensive text samples in a matter of seconds, making it a valuable tool for corpus linguistics, especially when studying rare syntactic structures, idiomatic expressions, or specialized registers. For example, a researcher investigating the diachronic use of subjunctive mood in conditional clauses could prompt ChatGPT to produce hundreds of sentences reflecting historical or regional usage patterns. Similarly, stylistic investigations into journalistic, academic, or literary tones can be simulated with ChatGPT by instructing it to generate texts in specific genres or voices.

However, one critical limitation is that ChatGPT is primarily trained to predict the most statistically probable next word based on a vast but finite dataset. As a result, it tends to reproduce dominant patterns—standard grammar, common collocations, and mainstream stylistic features. This may inadvertently reinforce language standardization, marginalizing less frequent or dialectal variations. For instance, attempts to simulate African American Vernacular English (AAVE) or regional Uzbek dialects may result in awkward or inaccurate approximations, as the model's training data likely underrepresents these varieties.

Moreover, in replicating "typical" usage, ChatGPT might oversaturate corpora with homogeneous structures, which can skew frequency-based analyses. If used uncritically, this could distort findings in studies aiming to model natural linguistic diversity (McGillivray, 2023). Therefore, while ChatGPT is a powerful tool for hypothesis testing and text generation, its outputs must be treated as artificial simulations, not authentic samples, and must be cross-validated with naturally occurring corpora. Methodological rigor, including metadata annotation and comparative analysis, is essential to prevent misleading conclusions.

3.3 Educational Technology and Language Pedagogy

Language educators increasingly use ChatGPT to generate tasks, grammar explanations, and vocabulary exercises. Beatty (2023) notes that this enhances

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 04, July, 2025
Website: usajournals.org

student engagement and access to materials. However, reliance on ChatGPT introduces risks of misinformation, grammatical errors, and diminished critical thinking. Therefore, teacher oversight remains essential to ensure pedagogical quality.

4. Linguistic Challenges Posed by ChatGPT

4.1 Bias and Representational Ethics

Because ChatGPT is trained on large-scale internet data, it often reflects societal biases—gender, racial, and cultural. Abid et al. (2021) document how LLMs exhibit anti-Muslim bias in sentence completions and associations. Bias may also appear subtly in how ChatGPT represents dialects, registers, or linguistic communities. Addressing this requires more than filtering: it demands ethical model design and inclusive data practices.

4.2 Lack of Grounding and Embodiment

ChatGPT lacks sensory experience—it cannot see, hear, or feel. Cognitive linguistics posits that language is deeply embodied and tied to perception and action (Lakoff & Johnson, 1980). Therefore, models without real-world grounding struggle to grasp spatial deixis, emotional nuance, or metaphoric concepts grounded in the body. This undermines their ability to truly understand language context.

4.3 Over-Reliance and Intellectual Ownership

Over-reliance on ChatGPT in writing or research may reduce originality and blur authorship lines. As AI-generated content becomes more common, determining who "owns" the language becomes complex. Floridi and Chiriatti (2020) argue that we must rethink authorship, agency, and accountability in the AI era. In linguistics, this also raises questions about using machine-generated data for theorizing about human language.

5. Philosophical and Methodological Implications

5.1 Redefining Linguistic Competence

Traditionally, competence referred to a speaker's internalized knowledge of language. ChatGPT complicates this, as it "performs" language well but lacks intentionality or consciousness. Is performance without competence still valuable

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 04, July, 2025
Website: usajournals.org

for linguistic theory? Many argue yes—for modeling and prediction—but it must be clearly differentiated from human cognition (Bender et al., 2021).

5.2 The Simulation–Understanding Divide

ChatGPT simulates understanding but does not experience it. This reflects a central concern in the philosophy of mind, most notably articulated by Searle's (1980) Chinese Room Argument. In this thought experiment, a person who does not know Chinese sits in a room and follows a set of rules to manipulate Chinese symbols. Although the person can produce coherent Chinese sentences that outsiders might interpret as meaningful conversation, the person has no actual understanding of the language—only the ability to manipulate form based on syntax. Similarly, ChatGPT processes and generates linguistic symbols without any grasp of their meaning.

For example, when prompted with a sentence like "Explain the difference between justice and revenge," ChatGPT may produce a fluent and logically structured answer:

"Justice is guided by law and seeks fairness, while revenge is driven by personal emotions and aims to inflict harm."

While this response is contextually appropriate and grammatically accurate, the model does not "know" what justice or revenge feel like, nor does it understand the social or moral consequences behind them. It simply draws upon statistical patterns from similar text strings in its training data.

Similarly, if asked to write a condolence letter, ChatGPT can output:

"I am deeply sorry for your loss. My thoughts are with you and your family during this difficult time."

While this appears empathetic, it lacks genuine emotional resonance because the model cannot experience grief, loss, or compassion. Its words are void of intentionality—there is no internal state, no emotional processing, no real "meaning" behind the language.

This distinction is crucial in linguistics and AI ethics. It means that although ChatGPT can mimic human discourse impressively, it does not engage in semantic interpretation or intent-driven communication. Its outputs are the result of probabilistic calculation, not conscious thought or comprehension.

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 04, July, 2025
Website: usajournals.org

Thus, while ChatGPT may pass surface-level linguistic tests like the Turing Test, it fails deeper tests of cognitive and experiential understanding. This raises important concerns about the limits of AI in domains requiring empathy, ethics, and human judgment, such as counseling, legal reasoning, or literary interpretation.

Conclusion

ChatGPT represents a major development in language technology with implications for nearly every subfield of linguistics. It challenges core assumptions about grammar, meaning, and communication while offering powerful tools for research and education. At the same time, it raises serious concerns about bias, grounding, and the limits of machine "understanding."

For linguists, the task is twofold: to embrace ChatGPT as a new instrument for exploring language, and to critically assess its limitations. This calls for a balanced approach that blends computational innovation with theoretical rigor and ethical awareness. ChatGPT forces us not only to reconsider what machines can do with language—but also what language really is.

References

1.Abid, A., Farooqi, M., & Zou, J. (2021). Persistent Anti-Muslim Bias in Large Language Models. arXiv preprint arXiv:2101.05783.
2.Beatty, K. (2023). AI in the Language Classroom: Promise and Pitfalls. TESOL Quarterly.
3.Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. Proceedings of ACL.
4.Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Proceedings of FAccT '21.
5.Berwick, R. C., & Chomsky, N. (2016). Why Only Us: Language and Evolution. MIT Press.
7.Chomsky, N. (1965). Aspects of the Theory of Syntax. MIT Press.
8.Floridi, L., & Chiriatti, M. (2020). GPT-3: Its Nature, Scope, Limits, and Consequences. Minds and Machines, 30(4), 681–694.

Modern American Journal of Linguistics, Education, and Pedagogy
ISSN (E): 3067-7874
Volume 01, Issue 04, July, 2025
Website: usajournals.org

9. Lakoff, G., & Johnson, M. (1980). Metaphors We Live By. University of Chicago Press.
10. McGillivray, B. (2023). The Role of LLMs in Modern Corpus Linguistics. Digital Scholarship in the Humanities.
11. Rogers, A., Kovaleva, O., & Rumshisky, A. (2020). A Primer in BERTology: What We Know About How BERT Works. Transactions of the ACL.
12. Searle, J. R. (1980). Minds, Brains, and Programs. Behavioral and Brain Sciences, 3(3), 417–457.
13. Vaswani, A., et al. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems.