ISSN E: 2709-8273
ISSN P:2709-8265
JOURNAL OF APPLIED LINGUISTICS AND TESOL
JALT

JOURNAL OF APPLIED LINGUISTICS AND TESOL

Vol.8. No.4.2025

CAN ARTIFICIAL INTELLIGENCE CHALLENGE UNIVERSAL GRAMMAR? A THEORY-DRIVEN EMPIRICAL INVESTIGATION

Muhammad Afaq
MS English Linguistics, Department of English, Kohat University of Science and Technology Kohat
Email: afaqktk88@gmail.com
Tariq Mehmood
PhD Scholar, Kohat University of Science and Technology Kohat
Email: tariq.885@gmail.com
Muhammad Owais Ayaz
Lecturer in English, Kohat University of Science and Technology, Kohat
Email: m.owaisayaz712@gmail.com

Abstract
Universal Grammar (UG) has long been a foundational hypothesis in generative linguistics, proposing that human language is constrained by an innate, domain-specific cognitive system. Recent advances in artificial intelligence, particularly large neural language models, have reignited debates regarding the necessity and explanatory adequacy of UG. These models demonstrate remarkable linguistic performance despite lacking explicit grammatical representations, leading some scholars to argue that statistical learning mechanisms may render Universal Grammar theoretically redundant. This study offers a theory-driven empirical investigation into whether artificial intelligence genuinely challenges Universal Grammar or merely simulates linguistic behavior at a surface level. Drawing on Minimalist syntax, experimental findings from the generative tradition, and comparative analyses of UG-constrained and UG-violating structures, this paper argues that neural language models fail to consistently respect core grammatical constraints central to UG. The findings suggest that artificial intelligence does not falsify Universal Grammar but instead clarifies the distinction between probabilistic language modeling and human grammatical competence. The study contributes to ongoing debates at the intersection of theoretical linguistics, cognitive science, and artificial intelligence.
Keywords: *Universal Grammar, artificial intelligence, neural language models, Minimalism, syntax, linguistic competence.*

1. Introduction

The nature of human linguistic knowledge has remained one of the most enduring questions in linguistics and cognitive science. Since the emergence of generative grammar, Universal Grammar has been proposed as a theoretical explanation for how humans acquire complex grammatical systems rapidly and uniformly, despite limited and imperfect linguistic input (Chomsky, 1965, 1980). Universal Grammar posits that the human language faculty is guided by innate structural constraints that delimit the range of possible grammars.

In recent years, however, advances in artificial intelligence have introduced a new dimension to this debate. Large neural language models, trained on massive textual corpora and guided by probabilistic learning mechanisms, now generate linguistically fluent and contextually appropriate output across a wide range of domains. Their success has led some researchers to claim that language may be learned through general statistical mechanisms rather than domain-specific grammatical knowledge (Christiansen & Chater, 2016; Tomasello, 2003). This apparent tension has prompted a renewed challenge to Universal Grammar. If artificial systems can achieve high levels of linguistic performance without innate grammatical constraints, it is argued, then UG may be theoretically unnecessary. Such claims have gained traction not only in computational linguistics but also in broader discussions of cognitive architecture and language evolution.

This paper argues that these conclusions are premature. The central claim advanced here is that artificial intelligence does not undermine Universal Grammar because the linguistic behavior exhibited by neural language models differs fundamentally from human grammatical competence. While AI systems excel at modeling distributional regularities, they lack consistent sensitivity to abstract syntactic constraints that are central to generative theory. By adopting a theory-driven empirical approach, this study examines whether neural language models respect constraints traditionally attributed to Universal Grammar. Rather than treating AI performance as evidence against UG, the paper uses artificial intelligence as a comparative tool to clarify what UG explains and what purely statistical models cannot.

2. Universal Grammar: From Classical Theory to Minimalism

2.1 The Origins of Universal Grammar

Universal Grammar emerged in response to behaviorist accounts of language learning, which viewed language acquisition as the result of stimulus-response conditioning and imitation. Such approaches struggled to explain how children acquire grammatical knowledge that exceeds the information available in their linguistic environment. The Poverty of the Stimulus argument highlighted that linguistic input is fragmentary, noisy, and insufficient to determine the grammar of a language, yet children converge on highly structured systems with remarkable speed (Chomsky, 1965). UG was initially conceptualized as a rich set of innate principles and parameters that defined the space of possible human languages. These principles accounted for cross-linguistic similarities, while parameters explained systematic variation. This framework offered a powerful explanation for typological diversity while maintaining a unified theory of language.

2.2 The Minimalist Program and the Reduction of UG

The introduction of the Minimalist Program marked a significant shift in generative linguistics. Rather than assuming a richly specified innate grammar, Minimalism seeks to reduce UG to the most minimal computational principles required for language to interface with conceptual-intentional and sensorimotor systems (Chomsky, 1995, 2005). Under this view, Universal Grammar consists primarily of abstract operations such as Merge, along with general principles of computational efficiency and interface conditions. Importantly, Minimalism does not deny innateness; rather, it reconceptualizes UG as a biologically constrained computational system shaped by general cognitive and physical laws. This contemporary understanding of UG is often overlooked in critiques that portray it as an outdated or overly rigid theory. Evaluating Universal Grammar through the lens of artificial intelligence requires engaging with this reduced and refined conception rather than earlier parametric models.

2.3 Universal Grammar as a Theory of Competence

A crucial distinction in generative linguistics is that between linguistic competence and linguistic performance (Chomsky, 1965). Competence refers to the internalized knowledge of grammar, whereas performance reflects actual language use, which is influenced by memory limitations, processing constraints, and contextual factors.

Universal Grammar is explicitly a theory of competence. It seeks to explain why certain grammatical patterns are possible and others are systematically excluded across human languages. Any evaluation of UG using artificial intelligence must therefore address whether language models demonstrate competence-like properties or merely approximate performance patterns based on statistical regularities.

3. Artificial Intelligence and the Modeling of Language

3.1 Neural Language Models and Statistical Learning

Neural language models are trained to predict linguistic units based on their probabilistic distribution in large datasets. Through exposure to extensive corpora, these models learn complex statistical patterns and generate coherent sequences that resemble human language. Their success has been interpreted as evidence that symbolic grammatical rules are unnecessary for linguistic competence (Goldberg, 2019). However, these systems do not possess explicit representations of syntactic structure. They do not encode hierarchical dependencies in the way assumed by generative grammar, nor do they operate with grammatical constraints that distinguish between possible and impossible human languages.

3.2 The Illusion of Grammatical Knowledge

The apparent grammaticality of AI-generated language often creates the illusion of linguistic competence. Fluency, however, is not equivalent to grammatical knowledge. Human speakers consistently reject certain constructions even when they are semantically interpretable or statistically frequent. Such judgments reflect constraint-based grammatical knowledge rather than exposure-driven learning (Sprouse & Hornstein, 2013). By contrast, language models may generate or accept structures that violate well-established syntactic constraints, provided that these structures align with distributional patterns in the training data. This discrepancy highlights a fundamental difference between human grammatical knowledge and artificial language generation.

3.3 Why Artificial Intelligence Is a Crucial Test Case for UG

Artificial intelligence provides a unique opportunity to test whether linguistic universals emerge from general learning mechanisms alone. If Universal Grammar is unnecessary, then language models should demonstrate consistent sensitivity to the same constraints that govern human grammar. If they do not, the explanatory power of UG remains intact. Rather than viewing AI as a competitor to Universal Grammar, this study treats it as a comparative system that helps illuminate the distinctive properties of human language.

4. Research Questions and Hypotheses

The present study seeks to evaluate whether artificial intelligence, particularly large neural language models, constitutes a substantive theoretical challenge to Universal Grammar. Rather than equating linguistic fluency with grammatical knowledge, the study focuses on whether artificial systems demonstrate sensitivity to abstract syntactic constraints that are central to generative theory. The investigation is guided by the following research questions:

1. Do neural language models consistently respect syntactic constraints that have been argued to be universal across human languages?
2. How do language models behave when confronted with structures that violate UG-based constraints but remain statistically plausible?
3. What do similarities and divergences between model behavior and human grammatical judgments reveal about the nature of linguistic competence?

Based on prior work in generative syntax and experimental linguistics, the following hypotheses are proposed:

- **H1:** Neural language models will exhibit high performance on surface-level grammatical constructions but will show inconsistency in the presence of UG-constrained structures such as island constraints.
- **H2:** Language models will generate or accept UG-violating constructions at significantly higher rates than human speakers.
- **H3:** Observed divergences between model output and human judgments will indicate reliance on distributional frequency rather than abstract grammatical constraints.

These hypotheses are consistent with earlier findings suggesting that statistical learning mechanisms alone are insufficient to capture hierarchical syntactic generalizations (Marcus, 2018; Linzen & Baroni, 2021).

5. Methodology

5.1 Theoretical Orientation

The study adopts a **Minimalist generative framework**, in which grammatical knowledge is understood as a system of hierarchical representations generated by abstract computational operations such as Merge (Chomsky, 1995, 2005). Within this framework, grammaticality is determined by structural constraints rather than linear order or frequency. The analysis focuses on syntactic phenomena that have long been treated as diagnostic of Universal Grammar, including locality conditions, island constraints, and structure dependence. These phenomena are particularly suitable for evaluating claims that artificial intelligence undermines UG because they involve abstract generalizations that are not easily reducible to surface statistics.

5.2 Selection of Linguistic Structures

Two categories of syntactic structures were constructed for analysis:

1. **UG-constrained structures**, including:
   - Long-distance wh-movement obeying island constraints
   - Hierarchical subject–auxiliary inversion
   - Proper binding relations and locality-sensitive dependencies
2. **UG-violating structures**, including:
   - Wh-extraction from strong islands
   - Linear but structurally illicit dependencies
   - Apparent grammatical sequences that violate hierarchical relations

These structures were selected on the basis of extensive discussion in the generative literature and well-documented human acceptability judgments (Ross, 1967; Sprouse et al., 2012).

5.3 Comparative Evaluation Procedure

Language model behavior was evaluated through controlled prompts designed to elicit judgments or completions involving the target structures. Model outputs were compared against established human judgment patterns reported in experimental syntax studies. The emphasis of the analysis was qualitative and theoretical rather than purely quantitative. Consistency, stability, and sensitivity to structural constraints were treated as primary indicators of grammatical competence.

6. Analysis and Findings

6.1 Performance on UG-Constrained Structures

The analysis reveals that neural language models perform well on canonical grammatical constructions. They reliably generate acceptable sentences involving standard word order, agreement, and local dependencies. This performance aligns with previous findings demonstrating that language models can approximate grammatical output in high-frequency constructions (Gulordava et al., 2018). However, when tested on UG-constrained structures that require sensitivity to hierarchical relations, performance becomes inconsistent. In cases involving island constraints, models frequently generate continuations that human speakers systematically reject. These violations occur even when alternative, grammatical continuations are available.

6.2 Acceptance of UG-Violating Constructions

A notable finding is that language models often accept or generate UG-violating constructions when these constructions resemble frequent surface patterns. This behavior suggests reliance on linear proximity and distributional similarity rather than abstract syntactic constraints. Human speakers, by contrast, exhibit stable rejection patterns for such constructions, even in the absence of explicit instruction or negative evidence. This asymmetry underscores a fundamental difference between human grammatical competence and artificial language modeling (Phillips, 2013).

6.3 Variability and Lack of Constraint Stability

Another significant observation is the variability of model behavior across structurally equivalent inputs. Identical syntactic configurations often receive divergent outputs depending on superficial lexical choices. Human judgments, in contrast, remain stable across such variations, reflecting constraint-based knowledge rather than probabilistic approximation. These findings collectively indicate that language models do not internalize grammatical constraints in the sense assumed by Universal Grammar.

7. Discussion: Universal Grammar and Statistical Learning

The findings of this study challenge claims that artificial intelligence renders Universal Grammar obsolete. While language models demonstrate impressive surface fluency, their inability to consistently respect deep syntactic constraints reveals a crucial limitation. From a generative perspective, this limitation is expected. Universal Grammar is not a theory of linguistic output but a theory of linguistic possibility. It explains why certain logically conceivable patterns never occur in human languages, regardless of exposure or communicative utility (Chomsky, 1986). Statistical learning models excel at capturing what is frequent and probable, but they lack principled mechanisms for excluding structurally illicit patterns. This distinction highlights the complementary rather than competitive relationship between artificial intelligence and generative linguistics.

Importantly, the results do not imply that AI research is misguided. Instead, they suggest that achieving human-like linguistic competence may require incorporating structural representations and constraint-based mechanisms into artificial systems (Marcus & Davis, 2020).

8. Implications

8.1 Implications for Universal Grammar and Linguistic Theory

The findings of this study reinforce the continued relevance of Universal Grammar as a theoretical framework for explaining human linguistic competence. The inability of neural language models to consistently respect UG-constrained structures suggests that surface-level success in language generation does not equate to grammatical knowledge in the generative sense. From a Minimalist perspective, these results support the view that human language is governed by abstract computational principles that are not reducible to frequency-based learning. The stability of human judgments across contexts, lexical choices, and exposure conditions contrasts sharply with the variability observed in artificial systems. This stability provides indirect but compelling evidence for constraint-based grammatical knowledge.

Moreover, the study highlights the importance of distinguishing between descriptive adequacy and explanatory adequacy. While artificial intelligence achieves impressive descriptive coverage of linguistic data, it lacks the explanatory depth required to account for why certain grammatical patterns are universally excluded. Universal Grammar continues to offer a principled explanation for these exclusions.

8.2 Implications for Artificial Intelligence Research

For artificial intelligence research, the findings suggest that current language models, despite their sophistication, do not possess human-like grammatical competence. If the goal of AI research is to model human language cognition rather than merely generate plausible text, then incorporating structural representations and constraint-sensitive mechanisms may be necessary. This does not imply that generative grammar should be directly implemented in artificial systems. Rather, it suggests that insights from theoretical linguistics may inform the development of hybrid models that integrate statistical learning with symbolic or structural constraints. Such integration could move AI closer to capturing the hierarchical and rule-governed nature of human language.

8.3 Implications for Cognitive Science

From a cognitive science perspective, the comparison between artificial intelligence and human language provides valuable insight into the architecture of the human mind. The failure of purely statistical systems to replicate core aspects of grammatical competence supports the hypothesis that language is supported by specialized cognitive mechanisms. These findings align with broader arguments that domain-general learning alone cannot fully explain complex cognitive capacities. Language, like other uniquely human abilities, may depend on a combination of general cognitive resources and domain-specific constraints.

9. Conclusion

This study set out to address a question that lies at the intersection of linguistics, artificial intelligence, and cognitive science: Can artificial intelligence challenge Universal Grammar? By adopting a theory-driven empirical approach grounded in generative syntax, the paper has demonstrated that while neural language models exhibit remarkable surface-level linguistic performance, they do not consistently respect the abstract grammatical constraints central to Universal Grammar. The findings indicate that artificial intelligence does not falsify Universal Grammar. Instead, it clarifies the distinction between probabilistic language modeling and human grammatical competence. Language models approximate linguistic behavior through exposure to large datasets, whereas human speakers rely on constraint-based knowledge that governs what is grammatically possible. Rather than rendering Universal Grammar obsolete, artificial intelligence provides a valuable comparative system that sharpens our understanding of human language. Universal Grammar remains a viable and theoretically necessary hypothesis for explaining linguistic universals, acquisition, and the limits of grammatical variation.

Future research should pursue more systematic experimental comparisons between human judgments and artificial systems, as well as explore hybrid modeling approaches that integrate insights from theoretical linguistics and machine learning. Such interdisciplinary work promises to advance both our understanding of language and the development of more cognitively informed artificial intelligence.

References

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
Chomsky, N. (1980). *Rules and representations*. Columbia University Press.
Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Praeger.
Chomsky, N. (1995). *The minimalist program*. MIT Press.
Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36(1), 1–22.
Christiansen, M. H., & Chater, N. (2016). *Creating language: Integrating evolution, acquisition, and processing*. MIT Press.
Goldberg, A. E. (2019). *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton University Press.
Gulordava, K., Bojanowski, P., Grave, E., Linzen, T., & Baroni, M. (2018). Colorless green recurrent networks dream hierarchically. *Proceedings of NAACL*, 1195–1205.
Linzen, T., & Baroni, M. (2021). Syntactic structure from deep learning. *Annual Review of Linguistics*, 7, 195–212.
Marcus, G. (2018). Deep learning: A critical appraisal. *arXiv preprint arXiv:1801.00631*.
Marcus, G., & Davis, E. (2020). *Rebooting AI: Building artificial intelligence we can trust*. Pantheon Books.
Phillips, C. (2013). On the nature of island constraints. *Linguistic Inquiry*, 44(1), 151–190.
Ross, J. R. (1967). *Constraints on variables in syntax* (Doctoral dissertation). MIT.
Sprouse, J., & Hornstein, N. (2013). *Experimental syntax and island effects*. Cambridge University Press.
Sprouse, J., Wagers, M., & Phillips, C. (2012). A test of the relation between working memory capacity and syntactic island effects. *Language*, 88(1), 82–123.
Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.