

Beyond the Hype: Feasibility and Conceptual Challenges of AI in Applied Linguistics Research

Nine Febrie Novitasari

Universitas Abdurachman Saleh Situbondo

ninefebrie@gmail.com

1. INTRODUCTION

The rapid growth of Artificial Intelligence (AI), particularly generative AI, over the past three years has created a new paradigm in language education globally (Smith et al., 2024). In this context, AI is no longer merely a passive background technology; rather, it exists as a dual instrument that fundamentally impacts the academic ecosystem by generating content and acting as a research object (Hetzscholdt, 2024). In other words, AI serves as a tool that assists researchers in analyzing data and compiling research results, while also positioning itself as a research object of interest to be examined from various pedagogical perspectives.

Currently, the dominant topics in Applied Linguistics literature focus heavily on ELT as the primary methodological practice. The majority of emerging studies focus on how AI affects students' language skills, such as writing and speaking (Al-nofaie & Alwerthan,

2024), affordances and challenges in its uses (Almegren et al., 2025), and effectiveness in the classroom (Hasbi, 2025; Rasool et al., 2025).

However, amidst this massive volume of research, there is a critical neglect of the theoretical foundations and methodological validity used. Despite this growing body of research, limited attention has been paid to the feasibility and conceptual validity of using AI as a subject or data source in linguistics research. Most studies tend to be descriptive and pragmatic, without questioning whether linguistic constructs designed for humans remain valid when applied to non-human entities (Dascal, 1992; Pinell, 2024) . This imbalance underscores the urgency of re-examining whether the current direction of AI research in ELT is based on a solid conceptual foundation or simply caught up in a momentary technological euphoria.

2. DISCUSSION

To address the previously raised doubts about conceptual validity, this discussion will outline the fundamental differences between using AI as a technical instrument and AI as an object of study. This sharpening of perspective is necessary to prevent researchers from falling into the misconception of equating machine statistical processes with human linguistic awareness.

2.1 AI as a Tool vs. AI as an Object in ELT Research: A Line of Demarcation

While AI in ELT is a practical evolution of older tools used for grammar checking and transcription, a conceptual "trap" emerges when researchers treat AI as a person rather than a piece of software. When studies analyze whether a chatbot is being "polite" or how well it "understands" social cues, they mistakenly attribute human-like intent and cognitive processes to what is actually just high-level statistical processing. Supporting the statement, Woudenberg et al., (2024) add that AI fails to meet the criteria of authorship because it cannot perform illocutionary speech acts such as promising or asserting, lacks the fitting mental states like knowledge, belief, or

intention, and cannot take responsibility for the texts it produces. In reality, AI does not choose to be nice or have a goal in a conversation; it simply uses complex math to predict word patterns based on massive amounts of data, enabling it to predict word patterns effectively (Cerf, 2023). This distinction is vital because, unlike humans who speak to share genuine feelings or ideas, AI generates text without any conscious awareness or "mind" behind the words (Negura, 2025). In other words, even though AI can reliably generate linguistic content that can be interpreted as "texts," they lack the linguistically mediated reality that language provides (Pinell, 2024), making their output a simulation of language use rather than genuine expression of feelings or ideas (Sambrotta, 2025). Failing to recognize this creates an illusion of understanding, leading researchers to flawed conclusions that treat mathematical outputs as if they were human thoughts. Ultimately, viewing AI as a "someone" rather than a "something" risks losing the essential human element of language, replacing genuine social interaction with programmed mimicry.

2.2 The Trend-Driven Research Phenomenon: Between Innovation and Superficiality

The academic world of ELT is currently experiencing what could be called "AI fever." Students and researchers often feel pressured to follow the latest trends to ensure their work is deemed relevant or easily published. This phenomenon has triggered a massive shift in research topics, sometimes ignoring the fundamental principles of theoretical fit and methodological feasibility. Many turn to AI research without a philosophical understanding of how large language models work (Kumar, 2024), which is crucial for effectively engaging with AI research and applications. As a result, there is a tendency to replicate older studies but replace human subjects with AI. For example, if students previously studied "Student Communication Strategies in a Speaking Class," they now try to study "AI Communication Strategies in Discussions." Pedagogically, this is a serious problem because it overlooks that human communication strategies arise from the need to overcome cognitive or sociolinguistic barriers (Baum, 2020) whereas AI "responses" stem

from mathematical optimization (Cerf, 2023). In addition to that, Lin (2025) critically identifies six interpretive fallacies that undermine the validity of replacing human participants with AI in behavioral research. This trend-driven research often sacrifices depth for speed, resulting in a limited understanding of genuine human-machine interactions (Guzman & Lewis, 2020) and futuristic findings that are actually conceptually fragile.

2.3 Conceptual Challenges and Methodological Feasibility: A Fundamental Critique

To understand why AI research in Applied Linguistics requires extra caution, we must dissect four key challenges that are often overlooked by today's researchers.

a. Lack of Human Agency

The essence of language is agency and intentionality (Searle, 1969). In Searle's Speech Acts theory, an utterance consists not only of words but also of the speaker's intention to do something (a promise, a command, a statement). AI, by its very nature, lacks intention (Montemayor, 2021). It simulates communicative behavior without participating in social meaning-making processes (Negura, 2025). When an AI "apologizes" for misinformation, it does not truly feel guilty or have a moral intention to improve social relations. This problem extends to the issue of identity construction. In Applied Linguistics, we study how language is used to construct personal, social, and cultural identities (Hasbi et al., 2025a). AI has no genuine life history, no body, and no emotions (Xin, 2018), which are crucial for authentic personal and cultural identity construction. Therefore, research attempting to analyze cultural identity in AI-generated text is actually analyzing training data bias, not a living identity construction (Pinell, 2024). Ignoring this lack of agency leads researchers to erroneous conclusions about the nature of the language they are studying.

b. Absence of Social Embeddedness

Language never exists in a vacuum, as it is always socially embedded (Bacchini, 2015). Theories like Brown and Levinson's

politeness (1987) are based on the concept of "face" and social risk in human interactions. People are polite because they care about long-term relationships and their social standing. AI has no social standing (Tække, 2025). It does not risk losing face and has no stake in power relations (Maas, 2023). When researchers attempt to analyze AI's politeness, they are essentially analyzing word distribution patterns that mimic the language etiquette of a data corpus. Without social embeddedness, linguistic analysis of AI output misses the most essential sociopragmatic dimension. Similarly, on the issue of gender, AI has no experience of being male or female, so the "gender language" it produces is merely a reflection of stereotypes embedded in big data, not an authentic expression of gender reflecting societal biases embedded in language, rather than the AI's own experience (Duan et al., 2025).

c. Absence of Developmental and Cognitive Processes

One of the pillars of ELT is an understanding of how humans acquire and learn languages (Basturkman, 1999). Theories such as interlanguage explain how students gradually construct linguistic systems through cognitive trial and error (Ellis, 1999). AI lacks this stage of cognitive development (Hasbi et al., 2025b). It does not make "mistakes" due to forgetfulness or native language interference like humans do. It makes errors due to parameter limitations or incomplete data (Rolls, 2024). Therefore, applying traditional error analysis to AI is conceptually inappropriate. AI does not possess pragmatic competence developed through real social interaction (Kecskés & Dinh, 2025). Researchers who fail to understand the distinction between human learning and machine training will produce highly misleading interpretations of language learners' abilities.

d. Prompt Dependency

From a methodological perspective, AI research faces significant challenges related to replication. Data in traditional qualitative and quantitative research typically possess a certain degree of stability (e.g., written text or recorded interviews) (Bernauer, 2015; Perrin,

2001). However, AI output is highly dependent on the prompt or instructions given (Hastomo et al., 2025; Yifeng, 2025). If a researcher evaluates the quality of AI-generated essays without documenting every detail of the prompt, temperature parameters, and model version, the data cannot be replicated. This instability makes many research claims about AI capabilities extremely fragile and only valid for a specific moment in time on a specific server (Antun et al., 2020).

2.4 Implications for Research Validity: The Threat of Overclaim

The inconsistency between human-centered theory and machine-centered research subjects ultimately threatens the validity of the research as a whole (Gebru et al., 2022). Many linguistic constructs were originally developed to explain human language use, making their direct application to AI-generated discourse theoretically problematic (Pinell, 2024). Construct validity is weakened when researchers no longer measure what they are supposed to measure. For example, if researchers measure AI creativity, are they measuring creativity in the human cognitive sense, or simply the algorithm's randomness? Without this clarity, AI-based ELT research risks becoming mere software testing wrapped in linguistic jargon, ultimately contributing minimally to our understanding of how humans learn and teach language.

3. FUTURE DIRECTION

After mapping out these challenges, this section offers a roadmap for researchers to avoid being trapped in criticism, but to be able to conduct more meaningful research.

Researchers must return to their theoretical roots. Do not treat AI as a second human. If they want to research AI, use an approach that recognizes its identity as a digital entity. Focus on analyzing text as a digital artifact or using AI as a catalyst to spark students' critical thinking. Researchers should be more courageous in exploring Critical AI Literacy, where the focus is not on how great AI is, but rather on

how humans respond to and critically analyze AI output (Long & Magerko, 2020).

The role of lecturers is not only to teach how to use AI, but also to guide the feasibility of research topics. Students need to be guided to develop feasibility awareness. They should ask themselves: "Is the theory I'm using really suitable for machines?" or "Does this research add value to my teaching practice?" Lecturers should encourage students to become more than just trend-followers, but also critical observers of technology, enhancing their understanding of technology's implications and encouraging innovative thinking in the field (Jones, 2022).

Future research should make greater use of mixed methods to capture the complexity of these interactions and draw on theory from the field of Human-Computer Interaction (HCI) to strengthen the theoretical foundation of Applied Linguistics. Moving beyond the AI hype requires academic courage to re-examine what we consider "language learning" and "language use." AI research in ELT must remain grounded in human values and theoretical rigor, so that technology does not overshadow our primary goal: understanding how language empowers people in an increasingly digitalized world.

Author



Nine Febrie Novitasari, S.Pd., M.Pd is a lecturer at Universitas Abdurachman Saleh Situbondo. She got her undergraduate and *Master's* degrees from Universitas Negeri Malang. She has academic experience in English Language Teaching and Applied Linguistics. Her research interests include media and material development, teacher professional development, and English for Young Learners. Besides teaching, she is actively involved in community service and academic publications. She can be reached through email at ninefebrie@gmail.com.

References

- Al-nofaie, H., & Alwerthan, T. A. (2024). Appreciative inquiry into implementing artificial intelligence for the development of language student teachers. *Sustainability*, 16(21), 9361. <https://doi.org/10.3390/su16219361>
- Almegren, A., Almegren, R. M., & Hazaea, A. N. (2025). AI powered ELT: Instructors' transformative roles and opportunities. *PLOS One*, 20(5), 1–10. <https://doi.org/10.1371/journal.pone.0324910>
- Antun, V., Renna, F., Poon, C., Adcock, B., & Hansen, A. C. (2020). On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proceedings of the National Academy of Sciences*, 117(48), 30088–30095. <https://doi.org/10.1073/pnas.1907377117>
- Bacchini, S. (2015). When language met society: Sociolinguistics in the twenty-first century. *English Today*, 31(2), 62–64. <https://doi.org/10.1017/S0266078415000139>
- Basturkman, H. (1999). A content analysis of ELT textbook blurbs: Reflections of theory-in-use. *RELC Journal*, 30(1), 18–38. <https://doi.org/10.1177/003368829903000102>
- Baum, S. D. (2020). Deep learning and the sociology of human-level artificial intelligence. *Metascience*, 29, 313–317. <https://doi.org/10.1007/s11016-020-00510-6>
- Bernauer, J. A. (2015). Opening the ears that science closed: Transforming qualitative data using oral coding. *The Qualitative Report*, 20(4), 406–415. <https://doi.org/10.46743/2160-3715/2015.2117>
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (4th ed.). Cambridge University Press.
- Cerf, V. G. (2023). Large language models. *Communications of the ACM*, 66(8), 15. <https://doi.org/10.1145/3606337>
- Dascal, M. (1992). Why does language matter to artificial intelligence? *Minds and Machines*, 2, 145–174.

<https://doi.org/10.1007/BF00704455>

- Duan, W., McNeese, N., & Li, L. (2025). Gender stereotypes toward non-gendered generative AI: The role of gendered expertise and gendered linguistic cues. *Proceedings of the ACM on Human-Computer Interaction*, 9(CSCW1), 1–35. <https://doi.org/10.1145/3701197>
- Ellis, R. (1999). Item versus system learning: Explaining free variation. *Applied Linguistics*, 20(4), 460–480. <https://doi.org/10.1093/applin/20.4.460>
- Gebru, B., Zeleke, L., Blankson, D., Nabil, M., Netehi, S., Homaifar, A., & Tunstel, E. (2022). A review on human–machine trust evaluation: Human-centric and machine-centric perspectives. *IEEE Transactions on Human-Machine Systems*, 52(5), 952–962. <https://doi.org/10.1109/THMS.2022.3144956>
- Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. *New Media & Society*, 22(1), 70–86. <https://doi.org/10.1177/1461444819858691>
- Hasbi, M., Nor, H., Madkur, A., Sulisty, T., Mulyani, Y. S. (2025). *Local-Based English Language Teaching Across Indonesia*. Rizquna.
- Hasbi, M., Islamiah, N., Sulisty, T., Marhaban, S., Sari, A. L., Nor, H. (2025). *Deep Learning: A Handbook for English Language Teachers*. Rizquna.
- Hasbi, M. (2025). The dual faces of deep learning in ELT: Meaningful and AI-driven approaches. In *Deep Learning: A Handbook for English Language Teachers* (pp. 1-11). Rizquna.
- Hastomo, T., Widiati, U., Ivone, F. M., Zen, E. L., Hasbi, M., & Khulel, B. (2025). AI-powered conversational agents and intercultural learning: Insights from Indonesian EFL students. *Intercultural Communication Education*, 8(1), 103217-103217. <https://doi.org/10.29140/ice.v8n1.103127>
- Hetzscholdt, P. (2024). Is AI giving us more than we can or even

- should handle? *Learned Publishing*, 37(1), 63–65.
<https://doi.org/10.1002/leap.1593>
- Jones, R. H. (2022). Commentary: Critical digital literacies as action, affinity, and affect. *TESOL Quarterly*, 56(3), 1074–1080.
<https://doi.org/10.1002/tesq.3153>
- Kecskés, I., & Dinh, H. (2025). “ChatGPT for intercultural pragmatic learning? Potentially, but not yet” – The question of using AI to develop students’ intercultural pragmatic competence. *Intercultural Pragmatics*, 22(2), 369–397.
<https://doi.org/10.1515/ip-2025-2008>
- Kumar, P. (2024). Large language models (LLMs): Survey, technical frameworks, and future challenges. *Artificial Intelligence Review*, 57(10), 1–51. <https://doi.org/10.1007/s10462-024-10888-y>
- Lin, Z. (2025). Six fallacies in substituting large language models for human participants. *Advances in Methods and Practices in Psychological Science*, 8(3).
<https://doi.org/10.1177/25152459251357566>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16.
<https://doi.org/10.1145/3313831.3376727>
- Maas, J. (2023). Machine learning and power relations. *AI & Society*, 38(4), 1493–1500. <https://doi.org/10.1007/s00146-022-01400-7>
- Montemayor, C. (2021). Language and intelligence. *Minds and Machines*, 31, 471–486. <https://doi.org/10.1007/s11023-021-09568-5>
- Negura, L. (2025). Simulated sense-making or social knowledge? Artificial intelligence and the boundaries of representation. *Journal for the Theory of Social Behaviour*, 55(3).
<https://doi.org/10.1111/jtsb.70012>
- Perrin, A. J. (2001). The CodeRead system: Using natural language

- processing to automate coding of qualitative data. *Social Science Computer Review*, 19(2), 213–220. <https://doi.org/10.1177/089443930101900207>
- Pinell, P. (2024). Does artificial intelligence speak our language?: A Gadamerian assessment of generative language models. *Political Research Quarterly*, 77(3), 713–728. <https://doi.org/10.1177/10659129241243038>
- Rasool, U., Li, J., & Liu, M. (2025). Perceptions of generative AI in teaching and learning: UTAUT and TPACK model-based study of teachers and students. *Acta Psychologica*, 261, 105827. <https://doi.org/10.1016/j.actpsy.2025.105827>
- Rolls, E. T. (2024). The memory systems of the human brain and generative artificial intelligence. *Heliyon*, 10(11), e31965. <https://doi.org/10.1016/j.heliyon.2024.e31965>
- Sambrotta, M. (2025). LLMs and the logical space of reasons. *Minds and Machines*, 35(4). <https://doi.org/10.1007/s11023-025-09751-y>
- Searle, J. A. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139173438>
- Smith, A. H., Higgs, J., Lizárraga, J. R., & Watson, V. W. M. (2024). Guest editorial: Artificial intelligence and composing just education futures. *English Teaching: Practice & Critique*, 23(1), 1–5. <https://doi.org/10.1108/ETPC-04-2024-202>
- Tække, J. (2025). Sociological perspectives on AI, intelligence and communication. *Systems Research and Behavioral Science*, 42, 574–584. <https://doi.org/10.1002/sres.3123>
- Woudenbergh, R. V., Ranalli, C., & Bracker, D. (2024). Authorship and ChatGPT: A conservative view. *Philosophy & Technology*, 37(1), 1–26. <https://doi.org/10.1007/s13347-024-00715-1>
- Xin, L. (2018). The Routledge handbook of language and identity. *Language and Intercultural Communication*, 18(3), 354–356. <https://doi.org/10.1080/14708477.2017.1314089>

Yifeng, Q. (2025). Prompt engineering in education: A systematic review of approaches and educational applications. *Journal of Educational Computing Research*, 63(7–8), 1782–1818. <https://doi.org/10.1177/07356331251365189>