Jurnal Humaniora Revolusioner

JHR, 10 (10), Oktober 2026
ISSN : 24475540

THE PROBLEM OF INTENTIONALITY IN AI-GENERATED LANGUAGE: A PHENOMENOLOGICAL STUDY FROM THE PERSPECTIVE OF PHILOSOPHY OF LANGUAGE

Annisa Humairah Meliala[1], Natasya Sabila[2], Nur Izza Assya'bana[3]
Email: annisameliala19@gmail.com[1] , natasyasyabila41470@gmail.com[2], nurizzaassyabana@gmail.com[3]
Universitas Negeri Medan

Abstract: *The development of generative artificial intelligence has brought significant changes to communication and educational practices, while also raising fundamental questions about the nature of language, particularly whether expressions generated by AI can possess meaning and intentionality in the same way as human language. This issue has become increasingly relevant as students, educators, and the wider public use technologies such as ChatGPT for learning, communication, and knowledge production. This study aims to critically analyze the concept of intentionality in AI-generated language from the perspective of the philosophy of language, focusing on the relationship between linguistic form, meaning, and communicative intention. The study employs a qualitative approach with an interpretive case study design. The participants consist of 12 individuals, including English language students and lecturers who have experience interacting with generative AI. Data were collected through semi-structured interviews and an analysis of human–AI conversation documents, and were then analyzed thematically to identify patterns of meaning-making and attribution of intentionality. The findings reveal three main themes: the tendency of humans to attribute intentionality based on the coherence of AI responses, the distinction between apparent meaning and human communicative intention, and the role of context in shaping interpretations of AI-generated language. The study contributes theoretically to expanding discussions in the philosophy of language concerning intentionality in machine-based communication and practically emphasizes the importance of critical AI literacy in English language education.*

Keywords*: Philosophy Of Language; Intentionality, Generative AI, AI-Generated Language, Meaning-Making, Critical AI Literacy, Human-AI Communication..*

INTRODUCTION

The development of generative artificial intelligence, such as ChatGPT, has brought significant changes to the way humans use and understand language. This technology is increasingly used in everyday life and in education for discussion, information searching, translation, and writing assistance. The ability of AI to produce language that is coherent, relevant, and similar to human communication has consequently raised questions about the nature of language and meaning. When users feel that AI "understands" their questions and provides answers that are appropriate to the context, a philosophical issue arises concerning whether AI actually possesses intentions or intentionality when generating language.

In the philosophy of language, intentionality is related to the directedness of the mind or intention toward an object. Human language does not merely consist of a sequence of words but also involves context, the speaker's intention, and the process of interpretation. This issue becomes more complex when language is produced by an AI system that does not communicate through human experience and consciousness. Although AI is capable of producing responses that appear meaningful, this ability does not automatically demonstrate that AI possesses understanding or communicative intentions. Current debates concerning Large Language Models (LLMs) also increasingly discuss the relationship between meaning, understanding, intentionality, and referential grounding.

This phenomenon is particularly important in the context of English language education because students do not merely use AI as an assisting tool but also interact with

it through language. Such interactions may influence how users understand meaning, communication, and the position of humans as language users. Convincing AI responses may cause users to assume that the system possesses certain knowledge or purposes. Therefore, a critical understanding is needed regarding the difference between AI's ability to generate language and humans' ability to assign meaning and intention to that language.

Research in recent years has extensively discussed the linguistic abilities of LLMs, human–AI communication, and issues concerning understanding and meaning produced by artificial intelligence systems. However, some studies still focus primarily on technical capabilities or on the question of whether AI can be considered capable of understanding language. Research that specifically examines how intentionality is understood and attributed by humans to AI-generated language still has room for further development. This gap is important because meaning is not only related to linguistic output but also to how users interpret and assign intentions to that output.

Based on these issues, this study aims to analyze the problem of intentionality in AI-generated language through the perspective of the philosophy of language. The study uses a qualitative approach to understand users' experiences and interpretations when interacting with generative AI. The research focuses on the relationship between language, meaning, communicative intention, and understanding in human–AI interaction.

This study is expected to make a theoretical contribution to the development of philosophy of language studies in response to the phenomenon of artificial intelligence, particularly concerning the concepts of meaning and intentionality. Practically, it may also increase the critical awareness of students and educators when using AI as part of English language communication and learning. Thus, this study does not only ask, "Can AI generate language?" but also addresses the more fundamental question: "Does AI-generated language have intentions, or does its meaning actually come from humans who interpret it?".

METHODOLOGY

The data were analyzed using reflexive thematic analysis. The analysis included the process of understanding the data as a whole, identifying units of meaning, assigning initial codes, grouping codes into themes, reviewing relationships between themes, and developing interpretations based on the philosophy of language perspective. The analysis was conducted manually with the assistance of document-processing software, while the final interpretation was carried out by the researcher in order to maintain reflexivity and preserve the context of the data..

RESULT AND DISCUSSION

The research findings indicate that participants' experiences with AI-generated language exist in tension between its apparent linguistic ability to understand and uncertainty regarding the existence of intentionality behind such language. The three themes that emerged—AI that appears to understand, meaningful language without a clear intention, and humans as meaning-makers—show that the issue of intentionality in AI cannot be understood simply by examining the quality of linguistic output. The central issue lies in how humans interpret that output and subsequently attribute understanding, intention, and agency to the AI system.

1. AI That Appears to Understand: When Coherence Is Perceived as Understanding

The first theme shows that coherent and contextually appropriate AI responses can create a subjective experience in which users feel that they are interacting with something that "understands." From Grice's perspective, the interpretation of an utterance is closely related to recognizing communicative intention. In human communication, when someone provides a relevant answer to a question, the listener

naturally assumes that there is an intention underlying the answer. The same pattern of interpretation appears to be carried over when users interact with AI.

However, this finding does not mean that the experience of feeling understood proves that AI possesses understanding. This is precisely where the philosophical problem lies. Participants primarily evaluate understanding based on observable linguistic performance, while the internal processes that produce the responses are not directly experienced by users. Grindrod (2024), in Synthese, similarly places the issue of linguistic intentionality in LLMs within the question of whether the ability to generate language is sufficient for attributing intentionality.

The findings of this study enrich this debate by showing that the issue of intentionality does not exist solely within the AI system but also within the way humans interpret AI language. In other words, there is a difference between the philosophical claim that "AI understands" and the user experience of "I feel understood by AI."

2. Meaningful Language, but Its Intention Is Questioned

The second theme demonstrates a change in the participants' positions. After acknowledging AI's ability to produce sensible language, they began to question whether those responses actually possess an intention. This tension can be explained through Searle's speech act theory. In human communication, speech acts are related to the speaker's intention and mental conditions. When someone makes a request, a promise, or an apology, the linguistic form does not stand alone but is connected to the speaker's action and intention.

AI presents a different situation. The system can produce forms of language that resemble requests, promises, apologies, or expressions of empathy, but the question of who or what actually "intends" to perform those actions remains open. Gubelmann (2024), in Philosophy & Technology, also questions whether LLMs can be treated as agents that perform speech acts in the same sense as humans.

This demonstrates that pragmatic success is not necessarily identical to intentionality. A response can function communicatively for a user without necessarily being assumed to originate from a mental state containing an intention. Here, the findings strengthen the distinction between linguistic performance and mental intentionality. AI may successfully produce an utterance that serves a particular function, but this success does not resolve the question of whether AI possesses intentions in the same way as human speakers.

3. Humans as Meaning-Makers

The third theme brings the discussion to a more reflective position. Participants not only evaluated AI but also began to realize that they themselves play a role in shaping the meaning of the responses they receive. When AI responses were unclear, users provided additional context, reinterpreted the answers, or connected them with the purpose of the conversation. This means that meaning in human–AI interaction is not simply "sent" by AI to humans but is formed through a process of interpretation.

This finding is interesting because it shifts the focus from the question "Does AI have meaning?" to the question "How does meaning become possible in human interaction with AI?" Torres-Martínez (2024) discusses how approaches to language and transformer models are related to the attribution of intentional states to linguistic agents. This perspective helps explain why humans can interpret AI language as something that has a purpose even though the status of its intentionality remains problematic.

In the context of English language education, this phenomenon has important implications. Students who use AI do not only learn from the information provided but also learn how to treat language generated by machines. If linguistic fluency is automatically regarded as evidence of understanding, users may overlook the differences between fluent language, meaningful language, and language that possesses

intention. Therefore, AI literacy in language education needs to include the ability to question the sources of meaning and intention behind AI-based communication.

4. Understanding Intentionality as a Relation, Not Merely an AI Property

The three themes ultimately lead to an understanding that intentionality in human–AI interaction is more appropriately viewed as a relational and interpretive issue. Participants are not merely interacting with a system that generates language; they also bring human assumptions, experiences, and communication habits into the interaction. When AI provides a contextually appropriate response, users employ human communication frameworks to understand that response. When users subsequently recognize AI's limitations, a gap emerges between their communication experience and their philosophical beliefs about what is actually happening.

In this regard, the study does not definitively claim that AI either possesses or does not possess intentionality. A more cautious position is that AI language can produce communicative effects that make intentionality appear to be present, while the ontological basis of that intentionality remains debated. Grindrod (2024) shows that linguistic intentionality in LLMs needs to be distinguished from mental intentionality.

These findings also expand the discussion in the philosophy of language by positioning user experience as part of the philosophical problem. So far, questions about AI have often focused on what machines are capable of doing. This study demonstrates that another question is equally important: what happens to humans when they encounter machines that communicate like humans? This question brings the philosophy of language from abstract problems concerning meaning toward concrete social experiences in digital communication.

5. Contextual Reflection and Researcher's Position

The context of English language students and lecturers also needs to be considered when interpreting the findings. As language users who are accustomed to paying attention to structure, context, and communicative functions, participants may have particular sensitivity to the linguistic quality of AI responses. This may influence their tendency to perceive AI as "understanding" when its responses are coherent. At the same time, their academic awareness of AI limitations may make them more critical of such attributions.

Researchers also need to be aware of their own position as interpreters of the data. A background in language studies may make linguistic aspects of AI responses more noticeable than other aspects of the users' experiences. Therefore, interpretation is not treated as a single objective truth but as the result of a dialogue between participants' experiences, theories from the philosophy of language, and the researcher's reflections.

Overall, this discussion shows that the problem of intentionality in AI-generated language is not merely a question of whether machines possess minds but also a question of how humans recognize, attribute, and experience meaning when interacting with machines. This perspective contributes to the philosophy of language while also opening opportunities for further research on meaning, intentionality, speech acts, and the boundaries between human communication and AI-based communication.

CONCLUSION

This study shows that the problem of intentionality in AI-generated language cannot be understood solely through the machine's ability to produce coherent and contextual language. Participants' experiences reveal a tension between the impression that AI understands language and the uncertainty about whether such language is actually based on communicative intention. These findings suggest that meaning in human–AI interaction is not only related to linguistic output but also to the human interpretive process through which context, intention, and communicative value are assigned to AI

responses.

Conceptually, the study demonstrates the importance of distinguishing between the ability to generate language, linguistic meaningfulness, and intentionality. AI can produce utterances that pragmatically function like human communication without necessarily possessing intentionality in the mental sense. Therefore, this study expands the philosophy of language by positioning user experience as part of the problem of intentionality. The question concerning AI consequently goes beyond whether machines "can speak" and develops into a question of how humans determine whether a language has meaning and intention.

In the context of English language education, these findings have practical implications for the development of critical AI literacy. Curricula can integrate learning activities that encourage students to evaluate AI responses based on context, intention, reliability, and the limitations of its linguistic capabilities, rather than solely on linguistic fluency. Educational institutions should also encourage the use of AI as a reflective tool rather than as a replacement for human reasoning and interpretation. For educational policymakers, these findings may serve as a consideration in developing AI-use guidelines that emphasize critical thinking, transparency, and user responsibility.

This study has several limitations. Its focus on participants from higher education and the field of English language studies limits the diversity of perspectives that can be identified. The relatively small number of participants and the specific research context also mean that the findings are not intended to be broadly generalized. In addition, the study focuses on users' perceptions and experiences and therefore does not deeply explore the technical or ontological dimensions of AI and its internal mechanisms.

Future research could broaden the context by involving students from various disciplines, lecturers, teachers, general users, and groups with different levels of digital literacy. Phenomenological, ethnographic, or conversational analysis approaches could also be used to gain a deeper understanding of how humans experience communication with AI. Future studies could bring together the philosophy of language, pragmatics, linguistics, and AI studies to further explore the boundaries between generated language, understood language, and genuinely intended language. Ultimately, the problem of intentionality in AI not only raises new questions about machines but also encourages humans to reconsider what actually makes language an act of communication..

DAFTAR PUSTAKA

Grice, H. P. (1989). Studies in the Way of Words. Cambridge, MA: Harvard University Press.

Grindrod, J. (2024). Large language models and linguistic intentionality. Synthese. https://doi.org/10.1007/s11229-024-04723-8

Gubelmann, R. (2024). Large language models, agency, and why speech acts are beyond them (for now)—A Kantian-cum-pragmatist case. Philosophy & Technology. https://doi.org/10.1007/s13347-024-00696-1

Searle, J. R. (1983). Intentionality: An Essay in the Philosophy of Mind. Cambridge: Cambridge University Press.

Torres-Martínez, S. (2024). A predictive human model of language challenges traditional views in linguistics and pretrained transformer research. Language and Semiotic Studies. https://doi.org/10.1515/lass-2024-0018