LANGUAGE, MEANING, AND INTENTIONALITY IN HUMAN–LLM INTERACTION: A PHENOMENOLOGICAL STUDY FROM THE PERSPECTIVE OF PHILOSOPHY OF LANGUAGE

[1]Theresia Tobing, [2]Cintia Marbun, [3]Laura Barus,[4]Crystalin Romian Manalu

[1,2,3,4]Universitas Negeri Medan
E-mail : reretbng94@gmail.com[1], linqhny@gmail.com[2], laurabarus6@gmail.com[3], cingyulbngaol@gmail.com[4],

ABSTRACT

The development of Large Language Models (LLMs) has raised fundamental questions concerning language, meaning, intentionality, context, and understanding. Although LLMs can generate coherent and contextually appropriate responses, their apparent communicative competence does not automatically establish that they understand language in the same way as human speakers. This study explores how users experience and interpret language generated by LLMs from the perspective of the philosophy of language. A qualitative phenomenological design was employed with ten LLM users from higher education institutions in Medan, Indonesia. Data were collected through semi-structured interviews and documentation of participants' interactions with LLMs. The data were analyzed thematically by focusing on meaning, intentionality, context, understanding, and communication. The findings indicate three central themes: LLMs as conversational partners that appear to understand, the distinction between linguistically accurate responses and contextual understanding, and the boundary between human language use and language simulation. The findings suggest that users may experience LLM responses as meaningful even when the status of machine intentionality and understanding remains philosophically contested. The study contributes to the philosophy of language by distinguishing between the possible internal status of understanding in an LLM and the meaning constructed by humans during human–AI interaction. It also highlights the importance of critical AI literacy in educational contexts.

Keywords : *Philosophy Of Language; Large Language Models; Intentionality; Meaning; Human–AI Interaction; Phenomenology*

INTRODUCTION

The development of generative artificial intelligence has transformed contemporary communication practices. Large Language Models (LLMs), including conversational AI systems, can generate written responses that are coherent, contextually related, and stylistically similar to human communication. These capabilities have increased the use of LLMs in education, translation, academic writing, information retrieval, and discussion. At the same time, they have raised fundamental questions about the nature of language, meaning, understanding, and intentionality.

A central philosophical problem concerns whether the ability to produce linguistically appropriate sentences is sufficient to establish that an artificial system genuinely uses language. Mahowald et al. (2024) distinguish between formal linguistic competence and functional linguistic competence. The former concerns the production of well-formed linguistic expressions, whereas the latter involves using language to achieve goals in the world. This distinction is important because the production of grammatically accurate and coherent text does not, by itself, settle the question of whether an LLM possesses understanding, communicative intention, or experience.

The relationship between LLMs and human communication has also been examined from a pragmatic perspective. Van Dijk et al. (2023) argue that debates about understanding and intentionality in LLMs require a nuanced approach rather than a simple opposition between genuine understanding and statistical imitation. From a pragmatic perspective, humans attribute mental states to others partly because such attributions help them interpret and predict behavior. This raises a further question: under what circumstances do users attribute understanding or intention to an LLM?

Recent research also demonstrates that the pragmatic abilities of LLMs require careful examination. Studies of pragmatic processing have investigated phenomena such as implicature, presupposition, reference, and deixis, while also identifying differences between model performance and human performance (Sravanthi et al., 2024). Such findings suggest that the evaluation of LLM language use should not be limited to grammatical fluency or surface coherence.

The philosophical problem is therefore not restricted to whether LLMs can produce language-like expressions. It also concerns the relationship between linguistic form, contextual meaning, communicative function, intentionality, and human interpretation. In ordinary communication, an utterance is interpreted in relation to context, speaker purpose, social relationships, and shared knowledge. Hall and Mazzarella (2023) emphasize the importance of pragmatic inference in understanding meanings that are not fully expressed by linguistic form alone.

Another relevant issue is the symbol-grounding problem. Harnad's work on symbol grounding has influenced discussions about whether a system can acquire meaning through symbolic relations without a sufficient connection to the world. In relation to LLMs, this question concerns whether the ability to produce sentences about experiences demonstrates that the system possesses those experiences. However, the philosophical debate remains open and cannot be resolved solely through the observation of fluent language generation.

Recent philosophical work continues to examine these questions. Van Dijk et al. (2023) provide a pragmatic perspective on understanding and intentionality, while Gubelmann (2024) discusses agency and speech acts in relation to LLMs. More recent scholarship also examines the

broader philosophical issues surrounding linguistic competence, meaning, communication, agency, and consciousness in language models (Millière & Buckner, 2026). These debates indicate that the relationship between language production and language use requires conceptual clarification rather than a simple binary answer.

Although previous research has examined the linguistic, pragmatic, and philosophical capabilities of LLMs, further attention is needed to the experiences through which users interpret apparent understanding during human–LLM interaction. A user may describe an LLM as understanding a question because its response is relevant and coherent, while simultaneously recognizing that the system may not possess human-like experience or intention. This experiential tension provides an important point of investigation for the philosophy of language.

Accordingly, this study examines how users experience and interpret the language generated by LLMs. It focuses on the distinction between linguistic accuracy and contextual understanding, the attribution of communicative intention, and the relationship between human interpretation and machine-generated meaning.

The study addresses the following research questions:

1. How do users experience and interpret the apparent understanding demonstrated by LLMs?
2. How do users distinguish between linguistic accuracy and contextual understanding in LLM-generated responses?
3. What philosophical implications do these experiences have for the concepts of language use, intentionality, meaning, and communication?

The study is expected to contribute to the philosophy of language by examining not only whether LLMs possess understanding but also how humans experience, interpret, and attribute meaning to machine-generated language. Practically, it may support the development of critical AI literacy among students and educators

Literature Review and Theoretical Framework

Language, Meaning, and Intentionality

Language is not simply a collection of grammatical forms. It is also used to express intentions, perform actions, establish relationships, and negotiate meaning in particular contexts. A pragmatic approach therefore examines language in relation to speakers, hearers, purposes, situations, and social expectations.

In human communication, the meaning of an utterance may depend on information that is not explicitly stated. For example, the expression “It is cold here” may function as a simple description, a request to close a window, or an indirect complaint, depending on the context. The linguistic form alone does not fully determine the communicative function.

This issue becomes more complex when the utterance is generated by an LLM. The system may produce a response that is appropriate to a user’s question, but it remains necessary to distinguish between the functional success of the response and the philosophical question of whether the system possesses an intention in the human sense.

Van Dijk et al. (2023) propose that understanding and intentionality should be examined with conceptual nuance. Their approach is useful for the present study because it avoids assuming that all apparently meaningful LLM outputs are either meaningless or equivalent to human communication.

In this study, intentionality is treated as a philosophical concept rather than as a directly observable variable. The participants’ statements provide evidence about how users perceive or attribute intention, but they do not independently establish the internal mental state of an LLM.

Symbol Grounding and the Problem of Meaning

The symbol-grounding problem concerns how symbols acquire meaning through their connection to the world. A system may manipulate symbols according to formal relationships, but the philosophical question is whether such manipulation is sufficient for genuine meaning.

In the context of LLMs, the issue arises because language models are trained on large quantities of data and generate responses by processing learned patterns. A model can produce sentences about rain, pain, loss, or happiness. However, the production of such sentences does not necessarily demonstrate that the model has personally experienced rain, pain, loss, or happiness.

For example, a user may ask an LLM to explain the experience of grief. The LLM may produce an emotionally appropriate response, and the user may find the response meaningful. Nevertheless, the meaningfulness of the response for the user should be distinguished from the claim that the system itself has experienced grief.

This distinction is central to the present research. The study examines how users interpret machine-generated responses and how they distinguish between the meaningfulness of an interaction and the possibility of human-like experience within the system.

Language as an Interactional and Situated Practice

A situated perspective views language as a practice shaped by social interaction, experience, context, and communicative purposes. Human language is acquired and used in relationships and activities that connect speakers to the physical and social world.

Beuls and Van Eecke (2024) discuss the importance of situated communicative interaction in human language learning and raise questions about how machine learning differs from human language acquisition. This perspective is relevant because LLMs and humans may produce similar linguistic forms through different processes.

A student interacting with an LLM may have a specific academic objective, prior knowledge, expectations, and personal experience. The meaning of the interaction is therefore not determined only by the generated response. It also emerges through the user's interpretation and practical use of the response.

This approach allows the study to examine human–LLM interaction as a communicative experience rather than treating language as a purely formal system.

Pragmatics and LLMs

Pragmatics focuses on language in context. In relation to LLMs, it includes questions about whether models can interpret implied meanings, context-dependent references, social expectations, and communicative intentions.

Research on pragmatic evaluation indicates that LLM performance varies across pragmatic phenomena. Sravanthi et al. (2024) developed a benchmark covering implicature, presupposition, reference, and deixis, and reported variation in model performance across these phenomena. Such research demonstrates that pragmatic competence cannot be assumed solely from fluent language generation.

Ma et al. (2025) similarly emphasize the importance of evaluating pragmatic abilities in LLMs through datasets and systematic assessment. Their work provides a broader context for understanding why apparent contextual appropriateness should be examined carefully rather than treated as conclusive evidence of human-like pragmatic understanding.

The present study differs from benchmark-based research because it focuses on users' experiences and interpretations rather than measuring model performance directly on standardized pragmatic tasks.

Theoretical Position of the Study

This study adopts a pragmatic perspective as its primary theoretical lens, supported by symbol grounding and situated interaction. The framework is organized around four analytical concepts:

1. Meaning: How users interpret the significance of LLM-generated responses.
2. Context: How users evaluate whether responses fit their situations and communicative needs.
3. Intentionality: How users attribute or question communicative purpose in LLM responses.
4. Understanding: How users distinguish apparent comprehension from human-like understanding.

The study does not assume that LLMs either possess or lack all forms of language use in an absolute sense. Instead, it examines the distinction between language generation, communicative functionality, human interpretation, and philosophical claims about understanding.

METHOD

Research Design

This study employed a qualitative phenomenological design. The design was selected because the research investigates users' experiences and interpretations when interacting with Large Language Models. Phenomenology is relevant to the study because it allows the researcher to examine how participants experience particular phenomena and how they assign meaning to those experiences.

The study focused on participants' perceptions of LLM-generated responses, particularly their experiences of apparent understanding, contextual appropriateness, communicative usefulness, and the difference between machine-generated language and human language.

The research did not attempt to determine the internal consciousness or mental status of LLMs through experimental testing. Instead, it investigated how users interpret the communicative behavior of LLMs and how those interpretations relate to philosophical concepts in the philosophy of language.

Research Participants

The study involved ten participants from higher education institutions in Medan, Indonesia. Participants were selected through purposive sampling based on the following criteria:

1. They were university students or active LLM users.
2. They were at least 18 years old.
3. They had experience using LLMs for communication or language-related tasks.
4. They were willing to participate in semi-structured interviews.

The original manuscript states that participant selection was based on information adequacy until data saturation was reached. However, the final manuscript should provide a more detailed explanation of how saturation or information adequacy was determined.

Information requiring confirmation before submission:

- Participants' age range.
- Academic background or field of study.
- Duration and frequency of LLM use.
- Main purposes of LLM use.
- Gender distribution, if ethically appropriate and relevant.

- Interview language.
- Date and location or platform of interviews.
- Whether ethical approval was obtained.

These details should be added only from the original research records.

Data Collection

Data were collected through semi-structured, in-depth interviews and documentation of participants' interactions with LLMs. Interviews were conducted either in person or online and lasted approximately 30–45 minutes.

The interview questions addressed the following areas:

- Participants' experiences using LLMs.
- Their perceptions of whether LLMs understand language.
- Situations in which LLM responses were appropriate or inappropriate to context.
- Their experiences of receiving linguistically accurate but contextually unsuitable responses.
- Their perceptions of the differences between human language and machine-generated language.
- Their views concerning meaning, intention, and communication.

Documentation of interactions was used to support the interpretation of interview responses. The documentation provided contextual evidence concerning the types of responses that participants described as meaningful, useful, appropriate, or inappropriate.

The final manuscript should explain whether the documented interactions were collected through screenshots, exported conversations, researcher notes, or another procedure. It should also clarify how personal information and potentially sensitive content were removed or protected.

Data Analysis

The interview data were analyzed using thematic analysis. The analytical process involved the following stages:

1. Repeated reading of interview transcripts and related documentation.
2. Identification of meaningful statements.
3. Open coding of relevant excerpts.
4. Grouping of similar codes.
5. Development of preliminary themes.
6. Review and refinement of themes.
7. Interpretation of the themes through the theoretical concepts of meaning, intentionality, context, understanding, and communication.

The analysis focused on participants' experiences rather than treating their statements as direct proof of the internal mental states of LLMs.

An example of the analytical structure is presented below. The table is an analytical model, and the final version should be completed using the actual coding records.

Data focus	Possible initial code	Analytical category	Philosophical relevance
A response appears to follow the user's question	Perceived comprehension	Apparent understanding	Attribution of understanding

A grammatically correct response does not fit the user's situation	Contextual mismatch	Accuracy versus understanding	Pragmatic appropriateness
The user values the usefulness of the response regardless of machine intention	Functional communication	Practical meaning	Instrumental interpretation of language
The user questions whether the LLM has personal experience	Lack of experience	Human–machine distinction	Symbol grounding and intentionality

The final manuscript should report the number of codes and themes only if these were documented during the actual analysis. NVivo or another qualitative analysis program may be mentioned only if it was genuinely used.

Trustworthiness

The study sought to support the trustworthiness of the findings through source triangulation, member checking, detailed contextual description, systematic documentation of research procedures, and the separation of empirical statements from researcher interpretations.

Source triangulation involved comparing interview findings with documentation of participants' interactions with LLMs. Member checking was intended to strengthen credibility by allowing participants to confirm or clarify the researcher's interpretation of their statements.

Dependability was supported through documentation of the research process, while confirmability was addressed by maintaining an audit trail and distinguishing participant data from the researcher's philosophical interpretation.

The final article should specify when and how member checking was conducted. It should also clarify whether participants reviewed transcripts, thematic summaries, or selected interpretations.

Ethical Considerations

Participants were informed about the research objectives, procedures, and use of research data before providing informed consent. Their identities were anonymized through codes or pseudonyms. Interview data and documentation of LLM interactions were stored securely and used for research purposes.

Participants were informed that they could refuse to answer particular questions or withdraw from the study without consequences.

The final manuscript should add the name of the relevant ethics committee and approval number if formal ethical approval was obtained. If no formal ethics approval was required or obtained, the authors should describe the institutional and ethical procedures that were followed accurately.

RESULT AND DISCUSSION

The thematic analysis identified three main themes concerning participants' experiences with Large Language Models:

1. LLMs as conversational partners that appear to understand.
2. The distinction between accurate responses and contextual understanding.
3. The boundary between human language use and language simulation.

These themes indicate that participants' experiences were not entirely uniform. Some participants emphasized the practical usefulness and communicative success of LLM responses, while others distinguished such usefulness from human-like understanding, experience, and intention.

LLMs as Conversational Partners That Appear to Understand

The first theme concerns participants' experiences of LLMs as more than simple technological tools. When participants used LLMs for discussion, writing, problem-solving, or information retrieval, the speed and apparent relevance of the responses sometimes created the impression of interaction with a conversational partner.

One participant stated:

"Sometimes I feel like I am talking to a person because the responses make sense and it can follow my questions." (P1)

This statement suggests that the participant evaluated the LLM's conversational performance through its ability to maintain relevance and respond to successive questions. The perception of conversational partnership was therefore associated with continuity, coherence, and responsiveness.

Another participant explained:

"I know that it is actually AI, but when the response matches what I mean, it feels like it really understands me." (P2)

This statement reveals a distinction between recognizing the system as artificial and experiencing its response as understanding. The participant did not necessarily claim that the LLM possessed human consciousness. Instead, the participant described a communicative experience in which the response appeared to correspond to the intended meaning.

From a pragmatic perspective, the experience of being understood may emerge through the success of an interaction. Users may attribute understanding to an LLM because such an attribution helps them describe and interpret the system's behavior. However, this experience should not be treated as conclusive evidence that the LLM possesses human-like intentionality.

The finding is consistent with the pragmatic perspective discussed by Van Dijk et al. (2023), which emphasizes that the attribution of understanding and intentionality requires conceptual nuance. In the present study, the participants' statements demonstrate that users may distinguish between what they know about the technology and what they experience during communication.

Accurate Responses Do Not Always Indicate Understanding

The second theme concerns situations in which participants received responses that were linguistically accurate but did not adequately address their intended meaning or context.

One participant stated:

"The answer is actually good, and the grammar is correct, but sometimes it is not what I mean. So, it is linguistically correct, but it does not really understand my context." (P3)

This statement explicitly distinguishes linguistic correctness from contextual understanding. The participant recognized that grammatical quality and apparent relevance do not necessarily guarantee that the response addresses the intended communicative purpose.

Another participant explained:

"I had to explain it several times to make the answer fit. From that, I felt that maybe it does not understand like a human does, but simply follows what I write." (P4)

The statement suggests that contextual difficulties became more visible when the interaction required personal knowledge, implicit information, or a specific communicative situation. The participant's interpretation was not that the LLM was incapable of producing language, but that its language production did not always correspond to the participant's intended meaning.

This theme can be interpreted through the distinction between linguistic form and communicative function. Linguistic form concerns the structure and apparent correctness of an utterance, while communicative function concerns what the utterance does in a particular situation.

The finding should not be interpreted as proof that all LLMs lack contextual understanding. Rather, it demonstrates that participants encountered situations in which linguistic fluency did not fully satisfy their communicative needs. It also shows why contextual appropriateness and pragmatic interpretation require separate consideration.

The distinction is relevant to research on pragmatic processing in LLMs. Studies of pragmatic benchmarks have reported variation across phenomena such as implicature, presupposition, reference, and deixis (Sravanthi et al., 2024). Such research supports the need to distinguish general linguistic fluency from performance on particular contextual and pragmatic tasks.

The Boundary Between Human Language and Language Simulation

The third theme concerns participants' reflections on whether the ability to produce useful responses is sufficient to classify an LLM as a language user.

One participant stated:

"If using language only means being able to form sentences and answer questions, then perhaps AI can do it. But if it requires having one's own intentions and experiences, I think it is different from humans." (P5)

This statement indicates that the participant distinguished between language production and language use understood as an activity involving intention and experience. The participant recognized that LLMs can produce sentences and respond to questions but questioned whether these abilities are equivalent to human language use.

Another participant expressed a more functional position:

"I do not really care whether it truly understands or not. What matters is that I can convey what I mean and get the response I need." (P6)

This statement suggests that communicative usefulness may be more important to some users than the ontological question of whether the system genuinely understands. The participant evaluates language through its practical success in achieving a communicative objective.

The two positions reveal a tension between an ontological interpretation and a functional interpretation of language. The ontological interpretation asks whether an LLM possesses understanding, intention, or experience. The functional interpretation focuses on whether the system's output enables users to communicate, solve problems, or complete tasks.

These positions should not be treated as mutually exclusive. A user may recognize the practical usefulness of an LLM while continuing to question whether the system possesses human-like understanding. This tension is central to the philosophical significance of human–LLM interaction.

Summary of Findings

The three themes indicate that participants experienced LLMs in an ambiguous space between technological tools and conversational partners. LLMs were perceived as

conversationally useful when their responses were coherent and relevant. However, participants also reported contextual limitations and questioned whether linguistic performance demonstrated genuine understanding.

The findings therefore suggest that the meaning of human–LLM interaction should be examined at two levels:

1. The meaning and usefulness experienced by the human user.
2. The unresolved philosophical question of whether the LLM itself possesses human-like understanding or intentionality.

The present study provides evidence about participants' experiences and interpretations. It does not independently establish the internal mental or ontological status of LLMs.

Discussion

Apparent Understanding and Human Interpretation

The first finding demonstrates that users may experience LLMs as conversational partners when responses appear coherent, relevant, and responsive. Participants' use of expressions such as understand, make sense, and follow my questions indicates that users employ ordinary human communicative concepts to describe interactions with artificial systems.

This finding can be interpreted through a pragmatic approach to mental-state attribution. In human communication, people commonly infer intentions and understanding from observable behavior. In interactions with LLMs, similar interpretive practices may be applied to machine-generated responses. Users may describe an LLM as understanding because its response appears to satisfy the communicative requirements of a particular exchange.

However, the experience of being understood and the possession of understanding are conceptually different. The first concerns the user's experience of the interaction, while the second concerns a philosophical claim about the nature of the system. The present data support the first claim but cannot independently settle the second.

This distinction is important because it prevents two forms of reductionism. First, it avoids reducing users' experiences to mere error or confusion. A user's experience of receiving a useful and meaningful response is a real communicative experience, even if the system does not possess human-like mental states. Second, it avoids treating conversational fluency as conclusive proof of human-like understanding.

The finding therefore supports an approach that examines both the technological output and the human interpretive process. Meaning in human–LLM interaction may emerge through the relationship between the generated response, the user's expectations, the communicative context, and the user's own knowledge and experience.

Linguistic Form and Communicative Function

The second finding indicates that participants distinguished between linguistically accurate responses and responses that adequately captured their intended meaning. This distinction is important for the philosophy of language because language is not merely a formal arrangement of words. It is also used to perform actions in specific contexts.

A response can be grammatical and informative but still fail to address the user's communicative purpose. For example, a user may ask a question involving personal circumstances, cultural assumptions, or implied meaning. An LLM may provide a well-formed response that does not fully account for those contextual factors.

This does not necessarily mean that the LLM cannot process context in any sense. Rather, it indicates that contextual appropriateness is multidimensional and depends on the task,

information available, interactional history, and standards used by the user to evaluate the response.

The distinction is consistent with research that treats pragmatic ability as a specific area of evaluation rather than as an automatic consequence of language fluency. Sravanthi et al. (2024) identified differences in LLM performance across pragmatic phenomena, while Ma et al. (2025) emphasized the need for systematic evaluation of pragmatic abilities in language models.

From a philosophical perspective, the findings support a distinction between:

- Formal linguistic performance: the production of grammatically and lexically appropriate expressions.
- Contextual appropriateness: the extent to which a response fits a particular situation.
- Communicative effectiveness: the extent to which the response helps the user achieve a communicative goal.
- Intentional understanding: the philosophical question of whether the system possesses intentions or understanding in a relevant sense.

These dimensions may be related, but they should not be treated as identical.

Symbol Grounding and Experience

The third finding concerns the relationship between language generation and experience. Participants questioned whether an LLM could be considered a language user if it lacked personal experience and intention in the human sense.

This issue is related to the symbol-grounding problem. A system may produce language about physical experiences, emotions, and social situations without necessarily undergoing those experiences itself. The ability to generate a description of an experience should therefore be distinguished from the possession of that experience.

At the same time, the absence of human-like experience in the system does not mean that the generated response is necessarily meaningless to the user. A person may interpret a machine-generated explanation through personal experience, cultural knowledge, and situational understanding. Meaning may therefore be constructed through the interaction between the user and the generated text.

This interpretation requires careful qualification. The study does not establish that all meaning exists solely in the user or that machine-generated outputs lack all forms of semantic structure. It demonstrates that participants' experiences of meaningfulness depend substantially on their own interpretive activity and communicative purposes.

The philosophical contribution of the study lies in separating the question of whether a system has understanding from the question of how its outputs become meaningful in human interaction.

Functional Communication and Ontological Understanding

The findings reveal two perspectives among participants. Some evaluated LLMs primarily through functional success: whether they could communicate an idea, provide a useful answer, or support the completion of a task. Others considered functional success insufficient for establishing that an LLM genuinely uses language in a human-like sense.

The functional perspective emphasizes the practical consequences of communication. If an LLM provides a useful response, the user may regard the interaction as successful regardless of whether the system possesses intentions. The ontological perspective, by contrast, asks whether the system has the properties traditionally associated with language users, such as experience, agency, intention, or understanding.

These perspectives should be distinguished rather than treated as competing empirical conclusions. They address different questions. Functional success concerns what the interaction accomplishes for the user, while ontological understanding concerns the nature and capacities of the system.

Gubelmann (2024) places questions of agency and speech acts at the center of the discussion about whether LLMs can be considered language users in a sense comparable to humans. This perspective is relevant to the present findings because participants distinguished between producing an answer and possessing one's own intention.

However, the current study does not directly analyze the technical architecture of LLMs or conduct a formal test of speech-act agency. Its contribution is to show how users encounter and interpret the distinction between useful language generation and human-like language use.

Implications for Philosophy of Language

The findings have three main implications for the philosophy of language.

First, they suggest that language use should not be analyzed solely through surface linguistic form. Users evaluate language through context, purpose, experience, and communicative effectiveness. This supports an approach that connects formal linguistic analysis with pragmatics and interaction.

Second, the findings show that meaning in human–LLM interaction may involve both machine-generated language and human interpretation. This does not resolve the question of machine meaning, but it demonstrates that the social and practical significance of LLM output cannot be understood independently of users.

Third, the findings support a distinction between the attribution of understanding and the demonstration of understanding. Users may attribute understanding to a system because its responses are useful and coherent, but such attribution does not necessarily establish the presence of human-like mental states.

The study therefore contributes to the debate by shifting attention from a single binary question—whether LLMs genuinely use language—to a set of related questions concerning linguistic performance, contextual appropriateness, human interpretation, communicative function, and intentionality.

Implications for Education and AI Literacy

The findings have implications for educational settings. Students and educators increasingly use LLMs for writing assistance, translation, discussion, and information retrieval. The usefulness of these systems creates opportunities, but it also requires users to evaluate their outputs critically.

AI literacy should therefore include more than technical knowledge about how to access and use LLMs. It should also involve understanding:

- The difference between fluency and accuracy.
- The difference between contextual relevance and genuine understanding.
- The possibility of incorrect or misleading responses.
- The importance of verifying information.
- The role of human judgment in interpreting generated content.
- The ethical and privacy implications of entering personal or confidential information into AI systems.

The educational implication of this study is not that LLMs should be rejected as communication tools. Rather, users should understand their capabilities and limitations and avoid assuming that fluent output is always equivalent to reliable knowledge or human understanding.

Limitations of the Study

This study has several limitations.

First, the number of participants was limited to ten users from higher education institutions in Medan. Their experiences cannot be assumed to represent all LLM users, including users from different regions, age groups, cultural backgrounds, professions, or educational contexts.

Second, the study focused on text-based interactions. It did not examine multimodal LLMs, voice-based interaction, visual input, or embodied systems.

Third, the study investigated users' experiences and interpretations rather than directly testing the internal processes of LLMs. Consequently, the findings cannot establish whether LLMs possess consciousness, human-like intentionality, or subjective experience.

Fourth, the interpretation of interview data may be influenced by the researcher's theoretical assumptions. Although triangulation, member checking, and documentation were used as reported in the methodology, the final article should provide more detailed evidence of how these procedures were implemented.

Future research could involve more diverse participant groups, comparative studies across educational and cultural contexts, digital ethnography, conversation analysis, and systematic examination of actual human–LLM interactions. Future studies may also combine phenomenological interviews with controlled pragmatic tasks to investigate the relationship between users' perceptions and observable model performance.

CONCLUSION

This study examined how users experience and interpret language, meaning, understanding, and intentionality during interactions with Large Language Models. The findings identified three central themes: LLMs as conversational partners that appear to understand, the distinction between linguistic accuracy and contextual understanding, and the boundary between human language use and language simulation.

The first finding indicates that users may experience LLMs as conversational partners when their responses are coherent, relevant, and responsive. However, the experience of being understood does not necessarily establish that an LLM possesses understanding in the same sense as a human speaker.

The second finding shows that linguistic accuracy and contextual understanding should be distinguished. A response may be grammatically correct and apparently informative while failing to address the user's intended meaning or specific communicative situation.

The third finding demonstrates that participants evaluated LLMs through both functional and ontological perspectives. Some participants emphasized the practical usefulness of communication, while others considered intention, experience, and understanding necessary for human-like language use.

From the perspective of the philosophy of language, the study suggests that human–LLM interaction should be examined at two interconnected levels. The first concerns the philosophical status of understanding, meaning, and intentionality within the system. The second concerns how users interpret and construct meaning through interaction with generated language.

The study does not establish whether LLMs possess human-like understanding or intentionality. Instead, it demonstrates how users negotiate these philosophical questions through practical communication experiences. This distinction provides a basis for further research on language, meaning, and communication in the era of artificial intelligence.

The study also highlights the importance of critical AI literacy in education. LLMs can support language-related activities, but users should critically evaluate their responses and distinguish linguistic fluency from reliability, contextual appropriateness, and human-like understanding.

REFERENCES

Beuls, K., & Van Eecke, P. (2024). Humans learn language from situated communicative interactions. What about machines? Computational Linguistics, 50(4), 1277–1311. https://doi.org/10.1162/coli_a_00534

Chen, X., Li, J., & Ye, Y. (2024). A feasibility study for the application of AI-generated conversations in pragmatic analysis. Journal of Pragmatics, 223, 14–30. https://doi.org/10.1016/j.pragma.2024.01.003

Dynel, M. (2023). Lessons in linguistics with ChatGPT: Metapragmatics, metacommunication, metadiscourse and metalanguage in human-AI interactions. Language & Communication, 93, 107–124. https://doi.org/10.1016/j.langcom.2023.09.002

Gubelmann, R. (2024). Large language models, agency, and why speech acts are beyond them (for now): A Kantian-cum-pragmatist case. Philosophy & Technology, 37, Article 32. https://doi.org/10.1007/s13347-024-00696-1

Gubelmann, R. (2024). Pragmatic norms are all you need: Why the symbol grounding problem does not apply to LLMs. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. https://doi.org/10.18653/v1/2024.emnlp-main.651

Hall, A., & Mazzarella, D. (2023). Pragmatic inference, levels of meaning and speaker accountability. Journal of Pragmatics, 205, 92–110. https://doi.org/10.1016/j.pragma.2022.12.007

Harnad, S. (2025). Language writ large: LLMs, ChatGPT, meaning, and understanding. Frontiers in Artificial Intelligence, 7, Article 1490698. https://doi.org/10.3389/frai.2024.1490698

Mahowald, K., Ivanova, A. A., Blank, I. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2024). Dissociating language and thought in large language models. Trends in Cognitive Sciences, 28(6), 517–540. https://doi.org/10.1016/j.tics.2024.01.011

Ma, B., Li, Y., Zhou, W., Gong, Z., Liu, Y. J., Jasinskaja, K., Friedrich, A., Hirschberg, J., Kreuter, F., & Plank, B. (2025). Pragmatics in the era of large language models: A survey on datasets, evaluation, opportunities and challenges. In Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (pp. 8679–8696). Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.acl-long.425

Millière, R., & Buckner, C. (2026). The philosophy of language models. Philosophy Compass, 21(3), Article e70095. https://doi.org/10.1111/phc3.70095

Ohmer, X., Bruni, E., & Hupkes, D. (2024). From form(s) to meaning: Probing the semantic depths of language models using multisense consistency. Computational Linguistics, 50(4), 1507–1556. https://doi.org/10.1162/coli_a_00529

Pezzulo, G., Parr, T., Cisek, P., Clark, A., & Friston, K. (2024). Generating meaning: Active inference and the scope and limits of passive AI. Trends in Cognitive Sciences, 28(2), 97–112. https://doi.org/10.1016/j.tics.2023.10.002

Rahman, A., & Liu, Z. (2023). The cognitive psychological distinctions between levels of meaning. Language Sciences, 100, Article 101583. https://doi.org/10.1016/j.langsci.2023.101583

Sravanthi, S., Doshi, M., Tankala, P., Murthy, R., Dabre, R., & Bhattacharyya, P. (2024). PUB: A pragmatics understanding benchmark for assessing LLMs' pragmatics capabilities. In Findings of the Association for Computational Linguistics: ACL 2024 (pp. 12075–12097). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-acl.719

Titus, L. M. (2024). Does ChatGPT have semantic understanding? A problem with the statistics-of-occurrence strategy. Cognitive Systems Research, 83, Article 101174. https://doi.org/10.1016/j.cogsys.2023.101174

Van Dijk, B., Kouwenhoven, T., Spruit, M., & Van Duijn, M. (2023). Large language models: The need for nuance in current debates and a pragmatic perspective on understanding. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (pp. 12641–12654). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.779