P-ISSN : 2654-5721
E-ISSN : 2654-7546
DOI : https://doi.org/10.37817/ikra-ithabdimas.v10i2

Who Speaks When AI Speaks? Agency, Meaning, and Responsibility in Human–AI Communication

Karjo Gantina Sinuhaji[1], Joan Graciela Silitonga[2], Meylan Keysia Abigail Sinaga[3], Najla Mutia Chalwa[4], Sola Gracia Marcelina Siregar[5], Christian Immanuel Samosir[6]

Abstract

The development of generative artificial intelligence (AI) in higher education has transformed the ways in which students interact with language, knowledge, and academic processes, while simultaneously raising questions about who speaks, how meaning is formed, and to whom agency and responsibility should be attributed when AI generates language. This study aims to understand students' experiences of human–AI communication, particularly in relation to the attribution of meaning, intention, agency, authorship, and responsibility to AI-generated outputs. The study employs a qualitative interpretive approach with a phenomenological design to explore the lived experiences of students who use generative AI in academic activities. Data are planned to be collected through semi-structured in-depth interviews, non-participant observation, and analysis of voluntarily provided AI conversation artifacts from 15–20 university students with experience using generative AI. The data will be analyzed using reflexive thematic analysis to identify patterns of meaning and experience emerging from human–AI interaction. The anticipated findings indicate four main patterns: AI is perceived as a non-human voice that can nevertheless be experienced as a conversational partner; AI's linguistic fluency can create an impression of understanding despite questions concerning intentionality; agency and authorship are negotiated through collaborative human–AI processes; and responsibility is understood as distributed and context-dependent. Theoretically, this study brings together Grice's theory of meaning, Speech Act Theory, and the perspective of relational agency to extend the study of language in the context of AI. Practically, the study emphasizes the importance of critical AI literacy in higher education so that the use of AI remains accompanied by human evaluation of meaning, reliability, agency, and responsibility.

Keywords: artificial intelligence, human–AI communication, meaning, agency, responsibility, philosophy of language, qualitative research

Theoretical Review

1. Language, Meaning, and the Problem of the "Speaker" in AI Communication

The emergence of generative artificial intelligence (AI), particularly large language models (LLMs), has brought fundamental questions in the philosophy of language back into focus: what makes an utterance meaningful, and is the ability to produce human-like language sufficient to describe a system as a "speaker"? LLMs are capable of generating coherent, contextual, and linguistically human-like responses. Consequently, users do not always treat AI outputs merely as

sequences of symbols but may experience them as responses from an entity that is engaged in communication.

This issue becomes more complex because the ability to generate language does not, by itself, demonstrate semantic understanding or intentionality. Bender et al. (2021), through the concept of "stochastic parrots," remind us that a model's ability to produce language resembling human language use cannot automatically be equated with an ability to understand meaning. This perspective is important for the present study because it distinguishes between similarity in communicative form and communicative status. In other words, AI may produce utterances that appear meaningful without necessarily being assumed to possess intentions in the same sense as human speakers.

Floridi and Chiriatti (2020) similarly demonstrate that GPT-3 is capable of generating text that closely resembles human language production while raising questions concerning understanding, intelligence, and the boundary between language simulation and cognitive ability. In higher education, these developments have become increasingly relevant because AI has entered learning processes, knowledge production, and academic activities. Dwivedi et al. (2023) demonstrate that generative AI has broad implications for society, organizations, education, and knowledge production.

Therefore, this study does not primarily begin with the question of whether AI "truly understands" language. Instead, it focuses on how students experience, understand, and interpret language generated by AI. This shift is consistent with the phenomenological orientation of the study, in which users' experiences become the point of entry for understanding how AI's status as a "speaker" is constructed in communicative practice.

2. Grice's Theory of Meaning: Meaning, Intention, and User Experience

H. P. Grice's theory of meaning serves as the primary theoretical foundation of this study because the concept of speaker meaning is directly related to the question, "Who speaks when AI speaks?" From a Gricean perspective, communicative meaning is not located solely in the linguistic form of a sentence but is related to what the speaker intends to communicate and how that intention is recognized by the listener.

Applying this theory to AI produces an intriguing problem. In human communication, when someone says, "I will help you," the listener can interpret the utterance based on context, social relationships, and the speaker's intention. However, when the same sentence is generated by AI, the question arises: who possesses the intention behind the utterance?

In this qualitative study, Grice's theory is not used to establish in advance whether AI possesses or lacks intention. Instead, it functions as a **sensitizing concept**, that is, a conceptual framework for examining how users themselves attribute meaning and intention to AI. For example, when a student states, "When ChatGPT explains something confidently, I feel that it really knows what it is talking about," the statement is not immediately understood as evidence that AI possesses knowledge. Rather, it is interpreted as an experiential account of how linguistic form, argumentative structure, and the degree of certainty expressed by AI create the impression that AI "knows" something.

P-ISSN : 2654-5721
E-ISSN : 2654-7546
DOI : https://doi.org/10.37817/ikra-ithabdimas.v10i2

Thus, Grice's theory enables this study to examine the process of attributing meaning rather than merely determining whether the information provided by AI is true or false. Even when an AI response is factually correct, users may still question whether AI genuinely "said" something or merely generated text that was subsequently interpreted by humans as an utterance.

This position is methodologically important because it protects the study from anthropomorphism. The researchers do not equate the user's experience that "AI is speaking to me" with the ontological claim that AI is a speaker in the human sense. Instead, the data consist precisely of users' experiences and language when they attribute a particular communicative status to AI.

3. Speech Act Theory: When AI Language Is Understood as Action

The second perspective is Speech Act Theory, which originated in the work of J. L. Austin and was subsequently developed by John Searle. This theory views language not merely as a means of describing reality but also as a form of social action. When a person requests, warns, promises, or states something, the utterance simultaneously constitutes part of a communicative action.

This theory is relevant when AI produces utterances that linguistically resemble human communicative actions. AI can provide recommendations, answer questions, make statements, or offer explanations. However, such linguistic similarity raises a fundamental question: is producing the linguistic form of an action equivalent to actually performing that action?

The issue becomes particularly evident in statements such as "I promise to help." In human communication, a promise has social and intentional dimensions and depends on certain conditions that enable a person to genuinely perform the act of promising. If such a sentence is generated by AI, its status as an action becomes problematic: is AI actually making a promise, or is it merely generating a sentence that linguistically takes the form of a promise?

In this study, this question is not directed solely toward determining the ontological status of AI. Speech Act Theory is used to examine how participants treat AI outputs as linguistic actions. One student may regard an AI recommendation as advice worthy of consideration, whereas another may perceive it merely as machine-generated output. Such differences indicate that the status of an AI utterance is also shaped through users' interpretive practices.

Accordingly, Speech Act Theory complements Grice's perspective. Whereas Grice helps the study understand how meaning and intention are attributed, Speech Act Theory helps explain what users perceive or understand AI to be doing through language—whether providing information, making a recommendation, issuing a warning, answering a question, or merely generating text.

4. Relational and Distributed Agency: AI as Part of the Communicative Process

The third approach shifts the question from "Does AI possess agency?" to "How is agency experienced and negotiated in the relationship between humans and AI?" This shift is important because agency

does not necessarily have to be understood as an internal characteristic of an entity. In human–AI interaction, users may attribute agency based on their interactional experiences. When AI provides relevant responses, maintains context, asks follow-up questions, or generates recommendations that were not explicitly requested, users may experience AI as something that is "doing" communicative actions.

Kasneci et al. (2023) demonstrate that the use of ChatGPT in education presents both opportunities and challenges for learning, teaching, and knowledge production. In academic practice, therefore, AI does not always function as a passive tool. AI can become part of the user's cognitive and communicative processes.

The concept of **distributed agency** helps explain this situation. An academic text, for example, may emerge through a sequence of actions: a student formulates a question, AI generates a response, the student selects particular sections, modifies them, verifies the information, and incorporates them into their writing. In such a process, it is difficult to identify a single point as the sole source of agency. Agency emerges across the user, the AI system, the instructions, the context, and the surrounding social practices.

This perspective prevents the study from adopting two extreme positions. The first considers AI to be an agent fully equivalent to a human being; the second considers AI merely a tool without social consequences. This study adopts a third position by understanding AI agency as a relational phenomenon that must be examined through users' experiences and interpretations.

This perspective is also relevant to understanding participants' emotional and social experiences. A student may feel helped, trustful, or comfortable when AI provides a response perceived as understanding their needs. However, this experience may change when AI provides inaccurate information. Such changes in trust demonstrate that the user's relationship with AI is not fixed but is continuously negotiated through interactional experiences.

5. Meaning, Agency, and Responsibility in Human–AI Communication

These three perspectives converge on the issue of **responsibility**. In human communication, speakers can generally be held accountable for what they say. In AI-mediated communication, however, this relationship becomes more complex.

For example, a student may obtain incorrect information from AI and subsequently include it in an academic assignment. The student may consider the information to have originated from AI. However, from the perspective of the educational institution, the student may still be regarded as responsible because the student chose to use and submit the information. At the same time, the AI output itself is the result of a system developed, trained, and provided through the involvement of various human and institutional actors. This situation indicates the possibility of **distributed responsibility**.

Importantly, this study does not determine in advance who should be responsible. Instead, it explores how participants themselves explain, justify, and negotiate responsibility. This approach allows contradictions to emerge as analytically meaningful phenomena. A student, for instance, may describe AI as "just a tool" when discussing inaccurate information while, in another situation, using highly personal language when asking AI for assistance. Such contradictions are not

P-ISSN : 2654-5721
E-ISSN : 2654-7546
DOI : https://doi.org/10.37817/ikra-ithabdimas.v10i2

regarded as weaknesses in the data but as important aspects of human–AI communicative experience.

6. Comparison and Theoretical Position of the Study

The three theoretical perspectives make different contributions to the study.

Approach	Main Question	Interpretive Focus
Gricean Theory	How is communicative meaning formed?	Meaning, intention, and user inference
Speech Act Theory	Is an AI utterance understood as an action?	Assertions, requests, advice, explanations, and linguistic actions
Relational/Distributed Agency	How is agency experienced and negotiated?	User–AI relationships, agency, authorship, and responsibility

Grice's theory is particularly strong in explaining the relationship between meaning, intention, and interpretation. However, the theory encounters limitations when the object of communication is not a human being. Speech Act Theory provides a more concrete framework for understanding language as action, but it still raises questions about who or what possesses the capacity to perform such actions. Meanwhile, the relational agency approach is more effective in explaining users' social experiences but may become too broad if it is not connected to questions of meaning and linguistic action.

Based on these considerations, this study positions **Grice's Theory of Meaning as the primary philosophical lens**, while Speech Act Theory and relational/distributed agency serve as supporting perspectives. This choice is based on the study's central focus on the word "speaks" in the question, *Who speaks when AI speaks?* The issue of "speaking" is fundamentally a question of meaning and intention, but meaning cannot be separated from linguistic actions or the social relationships within which such actions are interpreted.

Accordingly, the three perspectives are employed dialogically rather than as independent theories. The study does not seek to establish a priori whether AI "truly understands" language. Instead, it seeks to understand how humans experience AI as a speaking entity and what consequences this experience has for their understanding of agency, authorship, and responsibility.

7. Interpretive Conceptual Framework

The conceptual framework of this study is not constructed in terms of relationships among variables but rather as an interpretive map for understanding participants' experiences. AI is initially viewed as a producer of linguistic events experienced by users within particular contexts. The researchers then examine how users attribute meaning and intention to these responses using Grice's perspective as the primary analytical lens.

The study subsequently examines whether users understand AI responses as linguistic actions, such as advice, explanations, statements, or requests. The next stage traces how users attribute, reject, or negotiate agency in relation to AI. This

process is then connected to how users understand authorship and responsibility.

Conceptually, the analytical orientation can be represented as follows:

AI linguistic output → user experience → attribution of meaning/intention → interpretation as a speech act → attribution/negotiation of agency → attribution/negotiation of responsibility

However, this sequence is not understood as linear. Users' experiences may move back and forth between these dimensions. For example, receiving inaccurate information may alter a user's level of trust in AI, which may subsequently influence how the user understands AI's "intention" in later interactions.

Therefore, participants' voices are interpreted as experiences containing an ongoing interpretive process. When a participant states that "AI seems to understand me," the researcher does not convert the statement into a measurement of AI's level of understanding. Instead, the statement is interpreted as an experiential phenomenon by asking: what makes AI appear to understand, how is this experience formed, and how does this experience affect the user's relationship with AI?

Accordingly, the theoretical position of this study views the question *Who speaks when AI speaks?* not merely as a question of whether machines are capable of using language, but as a question of how speaker status, meaning, linguistic action, agency, and responsibility are negotiated within human communication with generative AI. This framework enables the study to retain philosophical questions concerning the nature of meaning and communication while also providing space for users' concrete social experiences with AI.

Results

The data analysis in this study is directed through reflexive thematic analysis to understand students' experiences when interacting with generative artificial intelligence (AI) in academic contexts. The focus of the analysis is not to objectively measure AI's capabilities but to understand how students experience, make sense of, and negotiate AI's status as an entity that "speaks."

Based on the experiential patterns addressed by the study, four main themes are identified: (1) AI as a non-human voice experienced as a conversational partner; (2) the impression of understanding produced by convincing language; (3) the negotiation of agency and authorship in human–AI collaboration; and (4) responsibility as distributed and context-dependent.

1. AI as a Non-Human Voice Experienced as a Conversational Partner

The first experience emerges when students use AI to support academic activities, such as seeking explanations, developing ideas, or discussing particular issues. In such situations, AI may initially be understood as a technological device. However, when interactions occur repeatedly and the responses appear relevant to previous questions, the boundary between a "tool" and a "conversational partner" becomes less clear.

The research manuscript describes this experience through users' expressions such as "AI told me..." or through experiences in which AI provides responses that feel appropriate to the context of the conversation. Such expressions are important not because they prove that AI is

P-ISSN : 2654-5721
E-ISSN : 2654-7546
DOI : https://doi.org/10.37817/ikra-ithabdimas.v10i2

a speaker in an ontological sense, but because they demonstrate how users attribute communicative status to AI.

For students, conversations with AI may feel different from merely entering commands into a technological device. AI can provide follow-up responses, maintain context, and respond dialogically to questions. These experiences may lead students to treat AI as something that is "responding" to them.

However, this experience is not always stable. A student may describe AI as a "discussion partner" in one situation but refer to it as a "tool" when discussing its limitations or errors. This shift indicates an important ambiguity: AI is technically not human, yet experientially it may appear as a communication partner.

Thus, this theme demonstrates that the status of a "speaker" is not determined solely by technological characteristics. It is also constructed through interactional experiences and users' interpretations of the responses they receive.

2. Convincing Language and the Impression that AI "Understands"

The second theme concerns situations in which students feel that AI understands the questions or problems they present. Responses that are well organized, argumentative, and expressed with a certain degree of certainty can create the impression that there is understanding behind the language.

In the research design, this experience is represented through the illustrative statement: "When ChatGPT explains something confidently, I feel that it really knows what it is talking about." This statement should be treated as an illustrative example within the research design rather than as an empirical quotation from an already interviewed participant.

Phenomenologically, the important question is not whether AI actually "knows," but how the feeling that AI "knows" is formed. Students may associate linguistic fluency with understanding. Fast, structured, and confident responses can make users feel that there is a subject that genuinely understands the issue being discussed.

Here, a tension emerges between the experience of understanding and the claim of actual understanding. Users may experience the interaction as meaningful communication while simultaneously recognizing that AI is not human. The research manuscript emphasizes that the ability to generate language resembling human communication does not, by itself, demonstrate understanding or intentionality.

This experience may also change when AI provides an incorrect answer. Information that previously seemed convincing may suddenly become questionable. Users may then question not only the content of the response but also whether the earlier impression of "understanding" can actually be trusted.

Thus, students' experiences reveal a paradox: the more convincing AI language becomes, the easier it is for users to experience AI as something that understands; yet subsequent errors may simultaneously reveal that such an impression is not equivalent to human understanding.

3. Agency and Authorship as a Process of Negotiation

The next experience emerges when AI is used not only to obtain information but also to assist in generating academic ideas or texts. In such practices, students do not necessarily accept AI outputs in their entirety. They may ask questions, receive responses, select particular sections, modify formulations, verify information, and subsequently integrate the material into their academic work.

The research manuscript describes this process as a sequence of actions between students and AI: students ask questions, AI generates responses, students select particular sections, revise them, and incorporate them into their writing.

This experience makes the issue of authorship less straightforward. Students retain the final decision regarding what is used, but some linguistic material and ideas may originate from interactions with AI. Consequently, agency is not always experienced as residing entirely with one party.

A tension arises when students attempt to determine "who actually created" a text. AI may be viewed as a source of assistance, while students remain the parties who select, edit, and use the resulting material. Conversely, when an error occurs, students may distance themselves from AI by describing it merely as a tool.

This contradiction is an important part of the experience. Agency in human–AI communication does not necessarily have to be understood as absolute ownership by either humans or AI. Rather, it can be understood as something that emerges and is negotiated through the relationship among users, systems, instructions, contexts, and academic practices.

Therefore, students' experiences indicate that the use of AI in academic activities is not simply a relationship in which "humans use a tool." It involves processes of selecting, interpreting, accepting, rejecting, and transforming AI outputs. These processes make agency and authorship continuously negotiable experiences.

4. Distributed and Contextual Responsibility

The issue of agency ultimately leads to the question of responsibility. When AI provides information that turns out to be inaccurate and that information is subsequently used in an academic assignment, a difficult question arises: who is responsible for what has been written?

The research manuscript indicates that students may regard AI as the source of an error, while educational institutions may continue to view students as responsible because students chose to use and submit the information. At the same time, AI outputs are produced by systems developed, trained, and provided through the involvement of various human and institutional actors.

This experience indicates that responsibility does not necessarily take the form of a single answer. In one context, a student may state that AI merely provides information and that the final decision remains with the user. In another context, the student may feel that the ability of AI to produce highly convincing answers makes the boundaries of responsibility increasingly difficult to determine.

This ambiguity becomes stronger when users treat AI differently depending on the situation. When AI helps them complete a task, users may speak to it personally and treat it as a partner. However, when an error occurs, AI may again be positioned as "merely a tool." The research manuscript explicitly treats such

P-ISSN : 2654-5721
E-ISSN : 2654-7546
DOI : https://doi.org/10.37817/ikra-ithabdimas.v10i2

contradictions not as weaknesses in the data but as important elements of the phenomenon requiring interpretation.

Therefore, responsibility in human–AI communication is understood as contextual and negotiated rather than as a category that can be determined simply before listening to users' experiences.

Thematic Synthesis

The four themes demonstrate a continuity in students' experiences with AI. The experience begins with AI responses as a linguistic voice, develops into an impression that AI appears to understand, proceeds into processes of collaboration and negotiation of agency, and ultimately raises questions concerning authorship and responsibility.

Main Theme	Experience Examined	Central Tension
AI as a conversational partner	AI is experienced as a responsive voice	Tool vs. communication partner
Impression of understanding	AI language feels convincing and contextual	Impression of understanding vs. intentionality
Negotiation of agency and authorship	Humans and AI participate in text production	Who "creates" or speaks?
Distributed responsibility	Users evaluate who is responsible for AI outputs	User responsibility vs. AI/system responsibility

These findings do not indicate a uniform experience. Rather, the meaning of AI as a "speaker" changes according to the context of interaction. AI may be understood as a tool, a conversational partner, a source of knowledge, or a component of text-production processes. Such shifts demonstrate that human–AI communication is dynamic and relational.

Overall, the findings suggest that the question *Who speaks when AI speaks?* cannot be answered simply by identifying AI or humans as the sole speaker. What emerges from users' experiences is a more complex process: humans interpret AI outputs as meaningful language, attribute a particular status to those utterances, respond to them, use them, and ultimately renegotiate agency and responsibility. Therefore, the "voice" of AI in students' experiences is not merely a question of who produces a sentence but also of who gives it meaning, who recognizes its communicative action, and who ultimately assumes responsibility for the use of that language.

Discussion

The findings indicate that students' experiences of communicating with generative artificial intelligence (AI) cannot be understood merely as a relationship between humans and a technological device. The four main patterns—AI as a non-human voice experienced as a conversational partner, the emergence of an impression that AI understands the user, the negotiation of agency and authorship in text production, and the understanding of responsibility as distributed and contextual—demonstrate that human–AI communication is an experience characterized by ambiguity.

These findings reinforce the theoretical position of the study that the question *Who speaks when AI speaks?* cannot be

adequately answered simply by determining whether AI is or is not a "speaker." The more fundamental questions concern how users attribute communicative status to AI, how meaning and intention are perceived, and how agency and responsibility are negotiated in communicative practice. Thus, "speaking" in the context of AI constitutes a phenomenon that is simultaneously linguistic, phenomenological, and social.

1. From Tool to "Voice": The Meaning of Experienced Communication

The first finding demonstrates that students may experience AI as more than merely a tool. Relevant responses, the ability to maintain context, and dialogic interaction patterns allow AI to appear as a "voice" in users' experiences. Nevertheless, this experience does not mean that students necessarily regard AI as human. Rather, there is continuous movement between two positions: AI is understood as technology while simultaneously being treated as a conversational partner.

This phenomenon can be explained through Grice's perspective on speaker meaning. In human communication, listeners generally connect an utterance to the speaker's intention. In interaction with AI, this relationship becomes unstable because users receive pragmatically convincing language without certainty that a human-like intention exists behind it. The research manuscript conceptualizes this as a distinction between intention phenomenologically experienced by the user and intention ontologically possessed by the system.

Accordingly, the finding should not necessarily be interpreted as evidence that AI has become a "speaker" but rather as evidence that speaker status can emerge through users' interpretive processes. When students say "AI told me," the expression reveals how users attribute a particular status to AI outputs. This position is consistent with Bender et al. (2021), who caution against equating linguistic similarity with human understanding.

However, this study also extends that discussion. Whereas the stochastic-parrot perspective primarily reminds us not to equate language production with understanding, students' experiences demonstrate that uncertainty regarding AI understanding does not eliminate the experience of communication. Users may still feel that they are communicating. Thus, the study understands meaning not merely as something that AI "possesses" but also as something experienced and interpreted by humans in interaction.

2. When Convincing Language Produces an Experience of Being "Understood"

The second finding reveals a relationship between the linguistic quality of AI responses and users' experiences that AI "understands" them. Coherent, systematic, and contextually appropriate responses may lead students to associate linguistic form with an underlying mental state. From a Gricean perspective, this demonstrates an inferential process: users encounter an utterance and then develop assumptions concerning the intention or communicative state underlying it.

Importantly, this study does not transform the experience "AI understands me" into a claim that AI actually possesses understanding. Instead, the experience is treated as a phenomenon requiring

P-ISSN : 2654-5721
E-ISSN : 2654-7546
DOI : https://doi.org/10.37817/ikra-ithabdimas.v10i2

interpretation. This position directly corresponds to the study's orientation, which treats users' statements regarding AI capabilities as experiential data rather than measurements of AI capability.

This finding can be considered alongside Floridi and Chiriatti (2020), who demonstrate that GPT-3's ability to generate language resembling human language production raises questions concerning the boundary between language simulation and cognitive ability. In education, Kasneci et al. (2023) likewise identify both opportunities and challenges associated with LLMs for learning and knowledge production.

The present study enriches this literature by highlighting an experiential dimension that may be overlooked when AI is discussed primarily in terms of technical capabilities. What is at stake is not only whether AI responses are accurate but also how trust, doubt, and the perception of being understood are constructed through language.

A significant paradox emerges here. More convincing language may strengthen users' trust, but subsequent errors may disrupt that experience. Therefore, the experience of being "understood" is not a fixed condition. It can be formed, reinforced, questioned, and even withdrawn through subsequent interactions. This finding indicates that trust in AI is a relational experience that develops through interactional history rather than merely as a consequence of the system's ability to generate fluent text.

3. Agency and Authorship: From Singular Ownership to Relational Processes

The third finding demonstrates that the issue of agency becomes increasingly complex when AI is used in academic writing. Students do not necessarily accept AI outputs in their entirety. They may formulate questions, select responses, edit them, verify information, and integrate them into their writing. Consequently, the final text emerges through a series of actions involving both humans and AI. The research manuscript describes this condition as agency distributed across users, systems, instructions, contexts, and social practices.

This finding reinforces the perspective of relational/distributed agency. Agency is no longer understood merely as an internal capacity possessed by a single actor but as something that emerges through relationships. Jarrahi et al. (2020), for example, describe human–AI relationships in knowledge management as a form of partnership that transforms work practices and decision-making.

In the present context, however, human–AI "collaboration" does not mean that humans and AI should be considered to possess identical forms of agency. Rather, they contribute differently within a sequence of practices. Humans determine goals, provide instructions, select, correct, and decide how outputs will ultimately be used, while AI provides linguistic outputs that are subsequently interpreted and processed.

This produces a complex problem of authorship. When students use AI to assist in generating text, the question "Who is the author?" does not always have an intuitive answer. However, the study suggests that the issue becomes more productive when understood as a negotiation of authorship rather than as a search for a single source of authorship.

Thus, an important contribution of this study is the shift from asking "Does AI possess agency?" toward the more phenomenological question: how do students experience their own agency

when part of the process of language production occurs together with AI? This shift is significant because AI use may alter students' experiences of writing, thinking, selecting, and making academic decisions.

4. Responsibility: When "AI Said It" Does Not Mean "AI Is Responsible"

The issue of agency ultimately leads to the question of responsibility. If AI is treated as a source of information or a conversational partner, users may attribute part of the responsibility to AI when an error occurs. At the same time, academic practice continues to position humans as the parties who use, verify, and submit the resulting material.

This finding demonstrates that responsibility in human–AI communication cannot be understood through a simple model that seeks a single responsible actor. The research manuscript emphasizes that responsibility may be understood as distributed and contextual, while deliberately avoiding a predetermined conclusion about who "should" be responsible.

Speech Act Theory provides an important contribution here. If an AI output is understood as a linguistic action—for example, advice, a statement, or a recommendation—then questions arise concerning who performs the action and who should bear its consequences. Austin and Searle demonstrate that linguistic actions are not determined merely by sentence form but also by the social and intentional conditions that make communicative actions possible.

In AI contexts, a tension therefore exists between **linguistic performance** and **communicative agency**. AI may generate linguistic forms resembling human actions without automatically acquiring the status of a human communicative agent. The findings demonstrate that users subsequently become the parties who interpret and make decisions based on those outputs.

Accordingly, responsibility in AI communication may be understood as a process that follows the trajectory of an utterance: from the system that generates the text, to the user who interprets it, and finally to the social or academic decision made on the basis of that text. This perspective expands the question of responsibility from "Who produced the sentence?" to "Who interpreted, used, authorized, and acted upon the sentence?"

5. Ambiguity as a Finding Rather Than a Weakness

One of the significant contributions of this study lies precisely in the ambiguity of students' experiences. Students may treat AI as both a tool and a partner, trust and doubt its responses simultaneously, and acknowledge its role in producing text while still retaining personal responsibility.

Such ambiguity should not be eliminated merely to produce a simpler conclusion. Instead, it reveals the actual character of human–AI communication. The research manuscript explicitly positions such contradictions as important elements of the data because they would remain invisible if the study relied exclusively on a normative approach.

From a phenomenological perspective, experience is not always logically consistent. An individual may simultaneously experience two apparently contradictory positions. In this context, students may rationally know that AI is not human while still experiencing AI as

P-ISSN : 2654-5721
E-ISSN : 2654-7546
DOI : https://doi.org/10.37817/ikra-ithabdimas.v10i2

something that “understands” or “speaks” to them.

The contradiction between **knowing that AI is a machine** and **experiencing AI as a conversational partner** therefore becomes an important characteristic of contemporary communicative experience.

This perspective enriches previous approaches that frequently place AI within the categories of “tool” or “agent.” The present study demonstrates that users’ experiences may occupy a position between these categories. AI may not be a human agent in the philosophical sense, yet it may become part of a relational configuration that influences how humans understand language, knowledge, authorship, and responsibility.

6. Theoretical Contribution and Contextual Reflection

Theoretically, this study extends the application of philosophy of language to AI-mediated communication by bringing together three perspectives: **Gricean Theory of Meaning, Speech Act Theory, and relational/distributed agency**.

The primary contribution is not to prove whether AI has become a speaker but to demonstrate that the category of “speaker” itself can become an object of negotiation. Grice helps explain how users attribute meaning and intention; Speech Act Theory demonstrates how AI outputs are treated as linguistic actions; and relational agency explains how agency and responsibility emerge through human–AI relationships.

In the context of university students and higher education, this issue becomes increasingly significant. AI is not merely a new technology but has become part of practices of reading, writing, learning, and knowledge production. Kasneci et al. (2023) and Zawacki-Richter et al. (2020) demonstrate the importance of understanding AI in relation to educational transformation and the role of humans within these processes.

Research reflexivity must also take the researchers’ positions into account. Because the study originates in philosophy of language and focuses on user experience, the researchers may bring particular assumptions regarding what constitutes “meaning,” a “speaker,” or “intention.” Therefore, participants’ statements should be treated as socially situated experiences rather than immediately being fitted into pre-existing philosophical categories. Grice’s theoretical position is used as an interpretive lens rather than as a tool for imposing an answer concerning whether AI possesses a mind or intention.

Within the sociocultural context of higher education, students’ experiences may also be influenced by academic norms concerning integrity, authorship, technology use, and responsibility for assignments. Therefore, when students state that AI “helps” or “provides answers,” such statements should be interpreted together with the institutional context in which AI use takes place. The meaning of an AI response does not exist independently but acquires consequences through the social practices surrounding it.

Ultimately, this study proposes that human–AI communication constitutes a space of meaning negotiation. AI generates language, humans provide interpretation, and the interaction subsequently produces consequences for agency, authorship, and responsibility. Thus, the question *Who speaks when AI speaks?* does not end with the answer that “AI speaks” or that “humans speak through AI.” Rather, the question opens a broader understanding that, in AI-mediated communication,

DOI : https://doi.org/10.37817/ikra-ithabdimas.v10i2
P-ISSN : 2654-5721
E-ISSN : 2654-7546

speaker status, meaning, action, and responsibility are continually reconstructed through the relationship among humans, technology, and social context.

This perspective constitutes a central contribution of the study: what changes through the presence of AI is not merely who produces language but also how humans determine what counts as meaning, who is considered a speaker, and when someone is considered responsible for an utterance.

Conclusion

This study demonstrates that human–AI communication in academic contexts cannot be understood merely as an interaction between a user and a technological tool. Four central patterns characterize the study: the experience of AI as a non-human voice that can be perceived as a conversational partner; the emergence of an impression that AI understands users through convincing linguistic outputs; the negotiation of agency and authorship in collaborative human–AI processes; and the understanding of responsibility as distributed and context-dependent. These patterns demonstrate that the meaning of communication with AI does not reside solely in system outputs but emerges through the relationship among AI language production, user interpretation, interactional context, and human decisions concerning the use of those outputs.

An important insight of this study is the tension between knowing that AI is not human and subjectively experiencing communication with it. Users may rationally understand that AI does not possess human-like experiences and intentions while simultaneously experiencing AI responses as communicative, relevant, and apparently understanding their intentions. Therefore, the question “Who speaks when AI speaks?” cannot be adequately answered by determining whether AI is a “speaker” or merely a “tool.” The deeper issue concerns how speaker status, meaning, agency, and responsibility are formed and negotiated within human–AI communicative practices.

Conceptually, this study reinforces the use of Grice’s theory of meaning as a perspective for understanding how users attribute communicative intentions to AI outputs without assuming that AI possesses intentions in the same way as humans. Speech Act Theory further helps explain how AI outputs may be treated as linguistic actions, while the perspective of relational agency demonstrates that agency is not necessarily attached exclusively to a single actor. In human–AI communication, agency and authorship may emerge through processes of selecting, interpreting, revising, and using AI outputs. Thus, the contribution of this study is not merely to position AI as an object of linguistic inquiry but to demonstrate the need to understand AI communication as a social-interpretive process involving humans, technology, and context simultaneously.

Practically, these findings have relevance for higher education and the development of AI literacy. Educational institutions should not only teach students how to use AI technically but also develop their ability to evaluate meaning, reliability, attributed intention, and the consequences of using AI-generated outputs. Digital literacy curricula may incorporate critical reflection on the differences between linguistic fluency and understanding, between convincing outputs and accurate information, and between AI assistance and academic decisions that remain the responsibility of humans.

For educational policymakers, this finding highlights the importance of AI-use

guidelines that focus not only on academic integrity but also on users' understanding of agency, authorship, and responsibility within human–AI collaboration. In the broader social context, such literacy is necessary to ensure that increasingly conversational interactions with AI do not automatically produce uncritical trust in technology.

This study also has limitations. Its focus on students' experiences in academic contexts limits the extent to which the findings can be generalized to other user groups, such as lecturers, researchers, professional workers, or the general public. Furthermore, relatively homogeneous participant characteristics may limit the exploration of variations in experience based on social and cultural backgrounds, academic disciplines, and levels of digital literacy. Limitations in time and depth of exploration may also constrain understanding of how perceptions of AI change after users encounter errors, inaccuracies, or repeated interactions. Therefore, understandings of agency and responsibility in human–AI communication remain open to further development.

Future research may broaden the context and population by involving students from different disciplines, lecturers, and AI users in diverse social and professional environments. Longitudinal approaches may also be employed to examine changes in the attribution of meaning, trust, agency, and responsibility throughout users' experiences with AI. Comparative research, digital ethnography, case studies, or mixed-methods approaches may provide broader perspectives on how culture, institutions, and technological characteristics influence experiences of human–AI communication. Further investigation into trust, authorship, AI errors, and the boundaries of human–AI responsibility will also be important for developing future research agendas in philosophy of language and communication studies in the age of artificial intelligence.

Ultimately, this study emphasizes that the central challenge of human–AI communication is not merely whether machines can produce human-like language, but how humans assign meaning to that language and determine who is considered to speak, act, and bear responsibility. This is where philosophical and practical concerns converge: the more fluently AI communicates, the more important it becomes for humans to maintain critical judgment concerning the meaning, agency, and responsibility that emerge within such interactions.

References

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. (2021). *On the opportunities and risks of foundation models*. Stanford Center for Research on Foundation Models. https://arxiv.org/abs/2108.07258

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ...

Wright, R. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management, 71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2020). A taxonomy of social cues for conversational agents. *International Journal of Human–Computer Studies, 132*, 1–15. https://doi.org/10.1016/j.ijhcs.2019.09.009

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines, 30*, 681–694. https://doi.org/10.1007/s11023-020-09548-1

Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2020). Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons, 63*(2), 163–175. https://doi.org/10.1016/j.bushor.2019.11.003

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), Article 248. https://doi.org/10.1145/3571730

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeiffer, F., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Luger, E., & Sellen, A. (2020). "Like having a really bad PA": The gulf between user expectation and experience of conversational agents. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3313831.3376638

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press.

Wilson, H. J., Daugherty, P. R., & Morini-Bianzino, N. (2020). The jobs that artificial intelligence will create. *MIT Sloan Management Review, 61*(4), 1–8.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2020). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education, 17*, Article 39. https://doi.org/10.1186/s41239-020-00191-9