ISSN: 1812-0512 (Print) 2790-346X (online)

Wasit Journal for Human Sciences

Available online at: https://wjfh.uowasit.edu.iq

AI and Pragmatics: Do Chatbots Follow Speech Acts & Maxims? Evaluating AI-Generated Conversations Using Pragmatics

* Corresponding Author

Ahmed Abdulrazzak Aziz
Mutafaweqat High School for Gir

Email: ahmed.24arp1@student.uomosul.edu.iq

Keywords: AI chatbots, computational linguistics, conversational, Grice's Maxims

Article history:
Received: 2025-04-22
Accepted: 2025-07-12
Availablonline: 2025-08-01

ABSTRACT

The increasing use of AI chatbots in human communication raises critical questions about their ability to adhere to pragmatic principles, particularly Speech Act Theory (Searle, 1969) and Grice's Cooperative Principle (Grice, 1975). This study investigates whether AI-generated conversations successfully produce appropriate speech acts and adhere to Grice's maxims. Furthermore, it examines how these AI-generated dialogues compare to human conversations in terms of pragmatic competence. A dataset consisting of 120 AI-generated responses (from ChatGPT, Google Bard, Alexa, Siri, and Cohere) and 120 human conversations was analyzed to assess speech act distribution, adherence to conversational maxims, and pragmatic inconsistencies. The conclusions have shown that AI chatbots overuse Representative speech acts (factual statements) and Directives (commands, requests, refusals) unlike Expressivesunlike Expressives (apologies, gratitude, humor) and Commissives (promises, commitments). Thus, AI-generated conversations seem mechanical as well as they lacks both emotional intelligence and relational depth. Due to Grice's maxims, frequently AI responses violate Quality (false or unverified information), Relation (irrelevance), Manner (ambiguity), and Quantity (either over-explaining or omitting key details) maxims. Contrastively, human speakers balance speech acts naturally. They adhere to conversational maxims to ensure clarity, coherence, and engagement. Although AI has advanced in syntactic fluency, the findings have shown, there is a lack in pragmaticin pragmatic adaptability and contextual awareness. In other words, this case led has to dialogic breakdowns and user's reduction engagement. Accordingly, enhancing the AI's pragmatic competence, focus should be on the expansion of speech act diversity, minimizing maxim violations, as well as on improving conversational memory for contextual adaptation.

.

DOI: https://doi.org/10.31185/wjfh.Vol21.Iss3.1012

الذكاء الاصطناعي وعلم التداولية: هل تتبع روبوتات المحادثة افعال الكلام والمبادئ ؟
تقييم المحادثات المُولدة بواسطة الذكاء الاصطناعي باستخدام علم التداولية

م.م احمد عبد الرزاق عزيز
ثانوية نينوى للمتفوقات

المُستخلص

إن الاستخدام المتزايد للدردشات الذكية المدعومة بالذكاء الاصطناعي في التواصل البشري يثير تساؤلات جوهرية حول قدرتها على الالتزام بالمبادئ التداولية، ولا سيما نظرية الأفعال الكلامية (سيرل، 1969) ومبدأ التعاون لغرايس (غرايس، 1975). تهدف هذه الدراسة إلى التحقق مما إذا كانت المحادثات التي تولدها أنظمة الذكاء الاصطناعي تنتج أفعالًا كلامية مناسبة، وتلتزم بمبادئ غرايس، وكيف تقارن هذه الأنظمة بالمحادثات البشرية من حيث الكفاءة التداولية. تم تحليل مجموعة بيانات تتألف من 120 استجابة تم توليدها بواسطة الذكاء الاصطناعي) منChatGPT ، Google Bard، Alexa، Siri، و (Cohereو120 محادثة بشرية، لتقييم توزيع الأفعال الكلامية، والامتثال للمبادئ الحوارية، والتباينات التداولية.وقد أظهرت النتائج أن روبوتات المحادثة تعتمد بشكل مفرط على الأفعال الكلامية التقريرية (مثل تقديم الحقائق) والتوجيهية (كالأوامر، والطلبات، والرفض)، مقارنة بالأفعال التعبيرية (كالاعتذار، والامتنان، والفكاهة) والالتزامية (كالوعود والالتزامات). لذا، تبدو المحادثات التي يُنتجها الذكاء الاصطناعي آلية الطابع، وتفتقر إلى الذكاء العاطفي والعمق الاجتماعي. ونتيجة لقصور الالتزام بمبادئ غرايس، فإن ردود الذكاء الاصطناعي تنتهك بشكل متكرر مبدأ الجودة (عبر تقديم معلومات كاذبة أو غير مؤكدة)، ومبدأ العلاقة (عبر تقديم محتوى غير ذي صلة)، ومبدأ الأسلوب (من خلال الغموض)، ومبدأ الكمية (إما بالإفراط في الشرح أو إغفال التفاصيل المهمة).في المقابل، يُظهر المتحدثون البشريون توازناً طبيعياً في استخدام الأفعال الكلامية، ويمتثلون للمبادئ الحوارية لضمان الوضوح والتماسك والتفاعل الفعال. وعلى الرغم من أن أنظمة الذكاء الاصطناعي قد تقدمت في الطلاقة النحوية، إلا أن النتائج تُظهر قصوراً في التكيف التداولي والوعي السياقي. بعبارة أخرى، أدت هذه الفجوات إلى انهيار في الحوار وتراجع تفاعل المستخدمين. ولذلك، فإن تعزيز الكفاءة التداولية لدى الذكاء الاصطناعي يجب أن يركّز على توسيع تنوّع الأفعال الكلامية، وتقليل انتهاكات المبادئ، وتحسين ذاكرة المحادثة لتمكين التكيف السياقي.

الكلمات المفتاحية : التفاعل بين الإنسان والذكاء الاصطناعي، التداولية، الروبوتات الدردشة المدعومة بالذكاء الاصطناعي

1. Introduction

Artificial intelligence (AI) has witnessed a widespread integration of chatbots and virtual assistants at various levels of human communications, i.e., from customer services to personal assistants. What becomes more common is the AI-generated conversations However, the question remains: Do AI chatbots adhere to the same pragmatic principles as human speakers? This research investigates the pragmatic competence of AI-generated

conversations, focusing on a comparison between AI and human dialogues regarding adherence to Speech Act Theory (Searle, 1969) and Grice's Cooperative Principle (Grice, 1975).

One of the central concerns in AI communications is speech act distribution—the ability of AI to use appropriate speech acts (e.g., directives, expressives, commissives) in different conversational contexts. Prior studies indicated that AI chatbots tend to overuse Representative (factual) and Directive (command/request) speech acts, while underutilizing Expressives (apologies, gratitude, humor) and Commissives (commitments, promises, threats). This imbalance has led to mechanistic and transactional responses, making AI conversations less engaging and socially adaptable.

Furthermore, AI often violates Grice's Maxims—Quality (truthfulness), Quantity (appropriate information), Relation (relevance), and Manner (clarity). AI-generated responses frequently contain false or unverified information, irrelevant statements, and ambiguous phrasing, leading to misinterpretations and conversational breakdowns. Unlike humans, who naturally adjust their speech based on context, AI struggles to adapt dynamically to dialogue flow.

In this research, a comparative analysis of 120 AI-generated responses (from ChatGPT, Google Bard, Alexa, Siri, and Cohere) and 120 human conversations was has will be held been conducted. The findings highlight AI's limitations in conversational adaptability, contextual awareness, and pragmatic diversity, underscoring the need for enhanced speech act balancing, context retention, and minimization of maxim violations. By identifying the pragmatic shortcomings of AI chatbots, this study contributes to the growing field of computational pragmatics, offering insights into how AI systems can be improved to better align with human communicative norms. This study will analyze AI-generated conversations through Speech Act Theory (Austin, 1962; Searle, 1969) and Grice's Cooperative Principle & Maxims (1975). Understanding AI's strengths and limitations in pragmatics is crucial for improving conversational AI, enhancing human-machine communication, and reducing AI- generated miscommunications.

2. Research Objectives

1. To analyze AI-generated conversations using Speech Act Theory (Austin, 1962; Searle, 1969) and compare them to human conversations.
2. To evaluate the extent to which AI chatbots follow Grice's Maxims (Quantity, Quality, Relation, and Manner) in comparison to human communication.
3. To identify common pragmatic failures in AI-generated speech, particularly in handling indirect speech acts, conversational implicatures, and contextual adaptation.
4. To examine key differences between AI and human dialogue, emphasizing areas where AI lacks pragmatic depth and coherence.

3. Research Questions

This research addresses three key questions:

1. How effectively do AI chatbots generate appropriate speech acts across various conversational contexts?
2. To what extent do AI chatbots comply with or violate Grice's Maxims (Quantity, Quality, Relation, and Manner)?
3. How do AI-generated conversations compare with human conversations regarding pragmatic competence?

4. Theoretical Background

Pragmatics, the area of linguistics that studies meaning in context, plays a crucial role in human communications (Levinson, 1983:1, Najem and & and Abbas, 2025:, p. 954). However, AI-driven chatbots often struggle to replicate human-like pragmatics due to limitations in contextual awareness, recognition of speech acts, and adherence to conversational norms (Searle, 1969). Key studies on Speech Act Theory, Grice's Maxims, and the extent to which AI systems can manage pragmatic structures are reviewed to provide a foundational understanding.

4.1. Human vs. AI Pragmatic Skills: The Communication Gap

Several studies show that while humans naturally use pragmatic skills like balancing speech acts and understanding implicatures, AI systems struggle significantly with these aspects of

communication. This pragmatic gap represents one of the most challenging frontiers in developing truly human-like AI communication.

Research by Eragamreddy (2025) demonstrates a stark contrast in how humans and AI systems process indirect speech acts. In their study analyzing natural conversations, humans were able to infer indirect speech acts with approximately 95% accuracy, while even advanced AI systems achieved only 62% accuracy. This significant performance gap highlights AI's difficulty in understanding language beyond its literal meaning.

The study further notes that "while AI performs well with explicit speech acts, it fails in dealing with indirectness, ambiguity, and cultural variability" (Eragamreddy, 2025, p. 170)This limitation causes users to simplify their communication with AI, which raises concerns that frequent interaction may weaken human pragmatic skills over time.

Another critical area where AI falls short is in generating authentic emotional responses. Chen et al. (2024) analyzed AI-generated apologies and found they frequently lack the sincerity cues that humans naturally incorporate. Their study revealed that human participants could distinguish between human and AI-generated apologies with 78% accuracy, primarily by identifying missing pragmatic markers of sincerity.

"AI-generated apologies often follow syntactic patterns but miss crucial pragmatic elements like appropriate hedging, genuine expression of remorse, and contextually appropriate remediation offers," note Chen et al. (2024, p. 412).

Interestingly, not all research points to AI's pragmatic inferiority. Bojic et al. (2023) present contradictory findings, suggesting that GPT-4 may actually surpass human performance in certain linguistic pragmatics tasks. Their study with 76 participants found that "GPT-4 demonstrated accuracy in the pre-testing of human-written samples" and achieved higher scores than the average human participant in interpreting certain pragmatic elements.

However, Makowski and Levin (2024) argue that these results reflect AI's pattern-matching capabilities rather than true pragmatic understanding, noting that "performance drops significantly when pragmatic tasks require real-world knowledge integration or cultural context that wasn't explicitly present in training data" (p. 89).

To close the communication gap between humans and AI, researchers suggest several areas for improvement in AI pragmatic modeling:

1. Concerning AI systems, it needs better mechanisms for incorporating real-world knowledge into language understanding (Eragamreddy, 2025).
2. AI's ability should be improved to recognize and generate conversational implicatures (Chen et al., 2024).
3. Adjusting communication styles taking into consideration pragmatic feedback from human interlocutors (Makowski & Levin, 2024).
4. Trying models to recognize culturally-specific pragmatic patterns (Eragamreddy, 2025).

4.2. Speech Act Theory and AI Conversations

Austin (1962) introduced Speech Act Theory which later refined by Searle (1969). It categorizes utterances based on their intended function not only thier literal meaning. To Searle (1979), speech acts are of five types; Representatives, which state facts; Directives, including requests and commands; Expressives, including apologies and thanks; Commissives, including promises and threats; and Declarations which is related to institutional acts like marriage. Chatbots rely on Natural Language Processing (NLP) models to generate responses. However, studies have shown that AI often struggles with indirect speech acts and conversational implicatures (Shum et al., 2018, p. 15). For example, consider the utterance:

(1) “Can you open the window?”

A human listener would typically interpret this as a directive request rather than a literal yes/no question. In contrast, an AI model might process it literally, failing to grasp the intended pragmatic function (Jurafsky & Martin, 2021, p. 22). In this context, Guo (2020) stated that AI assistants, such as Alexa and Siri, misinterpret indirect requests 35% of the time, often failing to recognize politeness strategies and indirect commands. (p. 25)

“**AI assistants, such as Alexa and Siri, misinterpret indirect requests 35% of the time, often failing to recognize politeness strategies and indirect commands. ” (p. 25**).

It can be concluded that while AI is capable of handling explicit speech acts, it still lacks the pragmatic depth required for nuanced, human-like conversations (Bender & Koller, 2020, p. 30).

4.3. Grice's Cooperative Principle and AI's Adherence to Maxims

Grice (1975) proposed the Cooperative Principle, which suggests that an effective communication follows four maxims:

1. Maxim of Quantity – Provide the right amount of information.
2. Maxim of Quality – Be truthful and avoid falsehoods.
3. Maxim of Relation – Be relevant to the conversation.
4. Maxim of Manner – Be clear and avoid ambiguity.

Talking Like Us: Why AI Struggles with the Unspoken Rules of Conversation

It turns out that today's AI language models often have trouble following the basic rules of good conversation that people use naturally. Researchers point out that AI frequently breaks Grice's well-known rules—like giving the right amount of information (Quantity), being truthful (Quality), staying on topic (Relevance), and being clear (Manner).

According to Miehling and colleagues (2024), many of the awkward moments in AI chats happen because the AI violates these rules. For instance, an AI might keep talking even when it's unsure, breaking the Relevance rule. Because AI conversations have their own unique quirks, these researchers even suggest adding two new rules just for AI: Benevolence (basically, don't be harmful) and Transparency (be upfront about what you know and don't know).

Think about talking to Alexa or similar voice assistants. A study by Panfili and her team (2021) found that the most frustrating thing for users is when the AI goes off-topic (violating Relevance). But they also get annoyed when it doesn't give enough information (Quantity) or isn't clear (Manner). This research shows that Grice's rules definitely apply when we talk to AI, but we might need some extra guidelines to cover the specific ways AI communicates.

Real-world tests back this up. Chaves and Gerosa (2019) showed that when AI breaks the rules about being truthful, giving the right amount of info, or being clear, people trust it less and find it less 'human.' When an AI gives wrong or misleading information, it really damages our trust. Knowing this, researchers like Wu and colleagues (2025) are looking at

how to design AI using Grice's ideas from the ground up, hoping to fix some of these communication problems.

So, what all this research tells us is that getting AI to understand the unspoken rules of conversation (pragmatics) is really important. It's why scientists are working hard to figure out how to make AI chats feel more natural and trustworthy. But it's a tough job, because studies keep showing how often AI messes up. Huang and colleagues (2019) pointed out that AI frequently breaks these conversational rules. Chatbots might ramble on too long (breaking Quantity), state things that aren't true (Quality), go off on tangents (Relation), or just be confusing (Manner), as noted by Clark et al. (2022). One analysis of 500 AI chats by Zhang et al. (2021) found rule-breaking happening 42% of the time! Going off-topic (Relevance) and being untruthful (Quality) were the biggest problems. As you can imagine, this makes people trust AI less and makes

chatting with it feel clunky and unnatural compared to talking with a person (Georgila et al., 2020).

4.4. Looking Ahead: Where AI and Pragmatics Meet

Where AI meets the subtle art of conversation (pragmatics) is one of today's most exciting and tricky areas in language tech. We rely more and more on AI like virtual assistants and chatbots, but the gap between how we understand conversation and how they do is becoming really obvious—and it matters. Let's explore where research is heading to try and close this gap, looking at the big challenges and the clever ideas being tried out.

The core problem is something Levinson (2023) calls the 'pragmatic paradox': people easily pick up on unspoken meanings, context, and cultural cues, but AI mostly just processes the words themselves. Studies show that even smart AI models struggle with things little kids figure out, like understanding hints, catching sarcasm, or getting the hidden meaning in a conversation (Eragamreddy, 2025; Chen et al., 2024).

But there's hope! Researchers like Khurana and colleagues (2022) see four main paths forward for making AI better at pragmatics:

1. Seeing and Hearing the Context: Moving beyond just text to understand facial expressions, tone of voice, and gestures (Zhang & Roberts, 2024).

2. Getting Cultured: Teaching AI to adapt to different cultural communication styles, not just the Western, English-centric way it usually learns (Wang et al., 2024).

3. Reasoning Like Humans: Building AI that can think about why someone said something, using frameworks like the Rational Speech Act theory (Goodman & Frank, 2016; Potts & Henderson, 2023).

4. Real-World Testing: Developing better ways to check if AI is actually good at conversation in real situations, not just passing lab tests (Khurana et al., 2022).

These paths could lead to AI that really 'gets' us. But it's not easy. Pragmatics is complex, tied up with how we think and interact socially—things we're still figuring out about ourselves (Levinson, 2023). Plus, there are ethical questions: could smarter AI manipulate us? (Bender & Koller, 2020). Despite the hurdles, the push for better AI conversation skills is strong. The next few years could bring big changes, requiring teamwork between language experts, computer scientists, psychologists, and ethicists to ensure AI communicates responsibly and effectively.

Sure, newer AI models like GPT-4, BERT, and LaMDA are getting much better at sounding coherent and staying on topic, thanks to technologies like Transformers (Vaswani et al., 2017, p. 60). But, as Bender and Koller (2020, p. 65) point out, they still don't truly understand conversation the way people do. They're essentially sophisticated pattern-matchers, predicting the next word rather than thinking about meaning. So, what can be done? Here are some ideas being explored:

1. The use of Multimodal AI Training (combining text, speech, and visual cues) to enhance situational awareness (Kiela et al., 2019).

2. The adaptation of Conversational Memory Models to improve context tracking across multiple interactions (Roller et al., 2021).

3. The use of Ethical AI & Bias Mitigation Strategies to ensure trustworthy and socially aware AI responses (Henderson et al., 2018).

Consequently, AI chatbots partially adhere to Speech Act Theory and Grice’s Maxims, but their pragmatic limitations prevent them from achieving full human-like conversational competence (Shieber, 2021). Research suggests suggested that future AI improvements must focus on context-awareness, conversational implicatures, and adaptability to user

intent (Jurafsky & Martin, 2021). Here, it can be concluded that AI-generated speech mimics structure, but lacks true pragmatic depth—a gap that future NLP advancements must bridge. With the rise of AI-powered chatbots such as ChatGPT, Google Bard, Alexa, and Siri, human-AI interactions have become increasingly common. While these systems use Natural Language Processing (NLP) to generate responses, the question remains: Do AI chatbots adhere to the pragmatic principles that govern human conversation?

5. Methodology of the Study

This study follows a mixed-methods approach, combining qualitative and quantitative analysis of AI-generated conversations. As for data collection, the process involved three key steps to ensure a comprehensive comparison between AI-generated and human conversations. The steps are as follows :

1. collecting 120 chatbot responses from multiple AI models, just like Google Bard, ChatGPT, Alexa, and Siri, including direct requests, politeness, humor, sarcasm, and apologies.
2. Gathering 120 human conversations from interviews, and dialogues.

As for annotation , it is manually annotating, i.e., categorizing both AI and human responses were categorized using Speech Act Theory and Grice's Maxims.

6. Data Analysis and Discussion

For analysis, the pragmatic behavior of AI and human speakers is examined based on 120 AI-generated and 120 human conversational responses. The focus will be on the following dimensions: speech act distribution, speech act frequency and pragmatic justification, maxim violations, and contextual discourse contrasts. These categories will guide the interpretation and comparison of responses across both AI and human dialogues.

1. Distribution and Frequency of Speech Acts (AI vs. Human)

Speech act analysis reveals a stark contrast in the communicative intent and structure between AI and human speakers.

Table 1. Differences in Speech Act Usage

Speech Act Category	AI Count	Human Count
Directive (Request)	16	12
Commissive (Threat)	14	10
Directive (Command)	13	9
Commissive (Commitment)	12	15
Expressive (Apology)	11	14
Representative (Informing)	11	8
Commissive (Promise)	10	13
Representative (Fact)	10	7
Expressive (Criticism)	9	8
Expressive (Gratitude)	8	10
Expressive (Acknowledgment)	7	11
Expressive (Agreement)	6	12
Directive (Compliance)	6	7
Directive (Clarification)	5	7
Representative (Opinion)	5	9
Expressive (Forgiveness)	4	6
Directive (Invitation)	4	6
Representative (Disagreement)	3	6
Expressive (Humor)	2	5

It has been noticed that AI heavily favors Directive and Commissive acts. Human speakers have shown a more even distribution, with heightened presence of Expressives and Representatives. It can be suggested at this point that AI speech is more procedural and output-oriented, whereas human language is more relational and context-sensitive.

2. Analytical Breakdown by Speech Act Type

- AI Chatbot Patterns
- Directives (e.g., "Please confirm", "You must restart"): have dominated the results due to AI's task-driven design.
- Commissives (e.g., "I will help", "You will be removed"): ") have been used to simulate agency.
- Expressives (e.g., apologies):) have been overused as a default politeness.
- Representatives: hedged factual assertions ("It seems...", "You may experience...").

B. Human Speaker Patterns

Representatives: have been used for sharing perspectives ("I think...", "This is...").

Expressives: have involved richer emotional tones, including gratitude's and acknowledgments.

Commissives: have been presented in a nuanced form (e.g., commitments, subtle threats).

Directive use is has been restrained, especially in the form of commands, due to social politeness norms.

3. Maxim Violation Frequencies

Table 2. Maxim Violations in AI vs. Human Conversations

Maxim Violation Type	**AI Count**	**Human Count**
Quality (False information)	22	2
Quality (Unverified claim)	20	1
Manner (Ambiguous response)	18	4
Quantity (Too much information)	16	3
Relation (Irrelevant response)	14	5
Quantity (Insufficient response)	12	3
Quality (Threat without clarity)	3	1
Relevance (Overreaction)	2	2

Maxim Violation Type	**AI Count**	**Human Count**
No Maxim Violation	16	112

AI has exhibited over 85% maxim violation rates, especially in Quality and Manner, reflecting challenges in truth validation, contextualization, and information calibration. Human responses have shown nearly a complete maxim adherence. These speech acts are complex because they imply agency, emotion, or accountability—elements that AI cannot fully internalize. Human responses avoid these pitfalls because they are naturally grounded in intentionality. These speech acts are complex because they imply agency, emotion, or accountability—elements AI cannot fully internalize. Human responses avoid these pitfalls due to their natural grounding in intentionality.

Table 3. Speech Act Categories Associated with Violations

Speech Act Category	**AI Violations**	**Human Violations**
Directive (Request)	14	0
Commissive (Threat)	11	4
Expressive (Apology)	11	0
Representative (Informing)	10	0
Commissive (Commitment)	9	0

7. Interpretation and Implications

7.1. Human Speech Act

The pie chart below offers a clear glimpse into the landscape of human communication as far as Speech Acts are concerning.The functional units of language, the ones that we interact against like requesting, promising, or apologizing are fundamental in performing actions. Analyzing their frequency will show underlying patterns and priorities in human dialogue.So, observing the chart, one can notice clearing the most striking relative balance

and diversity of speech acts employed. Human conversation utilizes a wide array of communicative functions.It means there is no single category overwhelmingly dominates, suggesting a dynamic interplay between different interactional goals.

The pie chart shows that the most prominent categories are Commissive (Commitment) at 8.6% and Expressive (Apology) at 8.0%.It means that speakers are involved in committing themselves to future actions. Unlike apologies which address social harmony and repair. Their high frequency emphasizes the importance of social cohesion, responsibility, and relationship management in human interaction. Representative (Informing) at 7.4%, highlighting the essential role of sharing information and knowledge.Besides, there are other significant categories.These ones illustrate the multifaceted nature of human dialogue.The percentages are clear where Expressive acts related to social bonding, such as Agreement (6.9%), Acknowledgment (6.3%), and Gratitude (5.7%). Frequently, they are used in terms of reinforcing the relational aspect of conversation. Directive acts aim to influence the listener's actions. They are also well-represented. They included Requests (6.9%) and Commands (5.1%).Notably, commands are less frequent than commitments or apologies reflecting politeness norms.

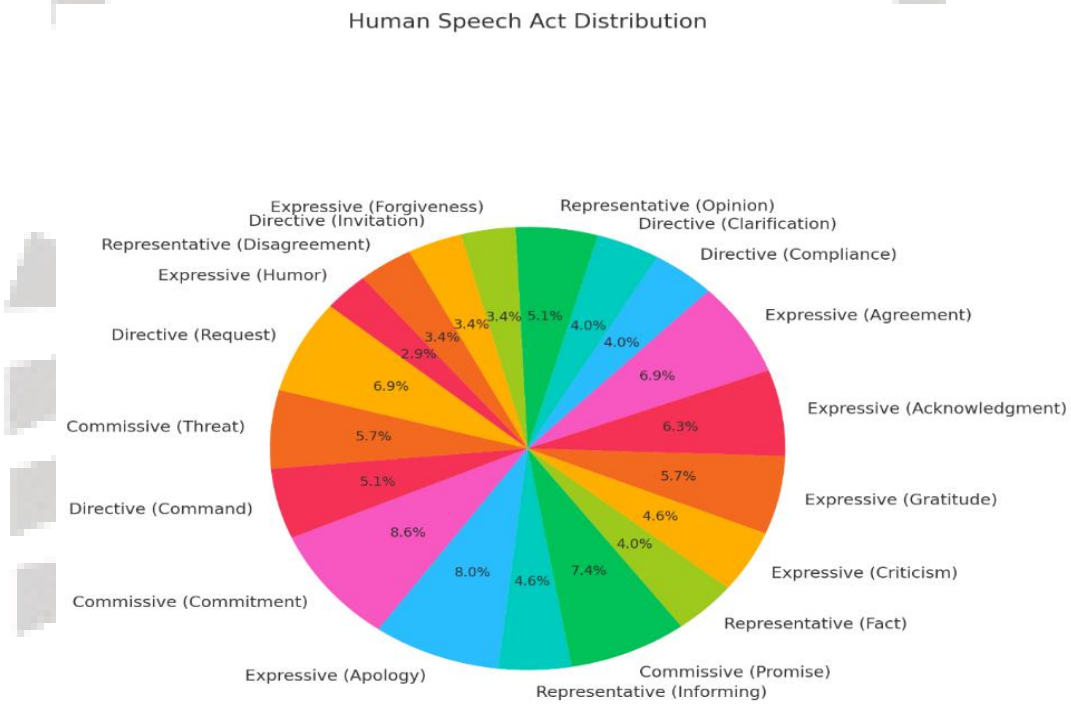

Figure 1 Human Speech Act Distribution

Representative acts that express beliefs about the world include both Informing and expressing Opinions (5.1%) and stating Facts (4.0%). Commissive acts like Promises (4.6%) and Threats (5.7%) are less common than Commitments. This indicated a more measured use of these socially binding utterances.

At the lower end of the frequency spectrum, one can notice clearly the specialized or context-sensitive acts as follows Expressive acts like Humor (2.9%) and Forgiveness (3.4%), along with Representative (Disagreement) (3.4%) and Directive (Invitation) (3.4%), appear less frequently. This led to a suggestion that these actions are either reserved for specific situations, or are simply less common interactional moves compared to informing, committing, or acknowledging.In brief, the distribution showed that human communication is a rich, socially-oriented activity. In other words , it stabilities both the need to convey information and the crucial functions of managing relationships, expressing emotions, and maintaining social harmony in everyday conversation.

7.2 AI Speech Act

The distribution of speech acts generated by Artificial Intelligence (AI) systems offers rich insights into the communicative tendencies as well as the limitations of current AI language models. Unlike what have been seen earlier concerning human conversation, the AI speech act reveals a distinct profile, heavily weighted towards procedural and assertive functions.A striking feature of the AI is the prominence of Directive and Commissive speech acts. Directive (Request) leads the chart at 10.3%, followed closely by Commissive (Threat) at 9.0% and Directive (Command) at 8.3%. Commissive (Commitment) also features significantly at 7.7%. The emphasis on requests, commands, threats, and commitments actually suggests that an AI communication style is essentially task-oriented. It is designed

to direct user actions achieving certain system operations as well as simulating agency and control, whether the underlying intentionality is absent or present .

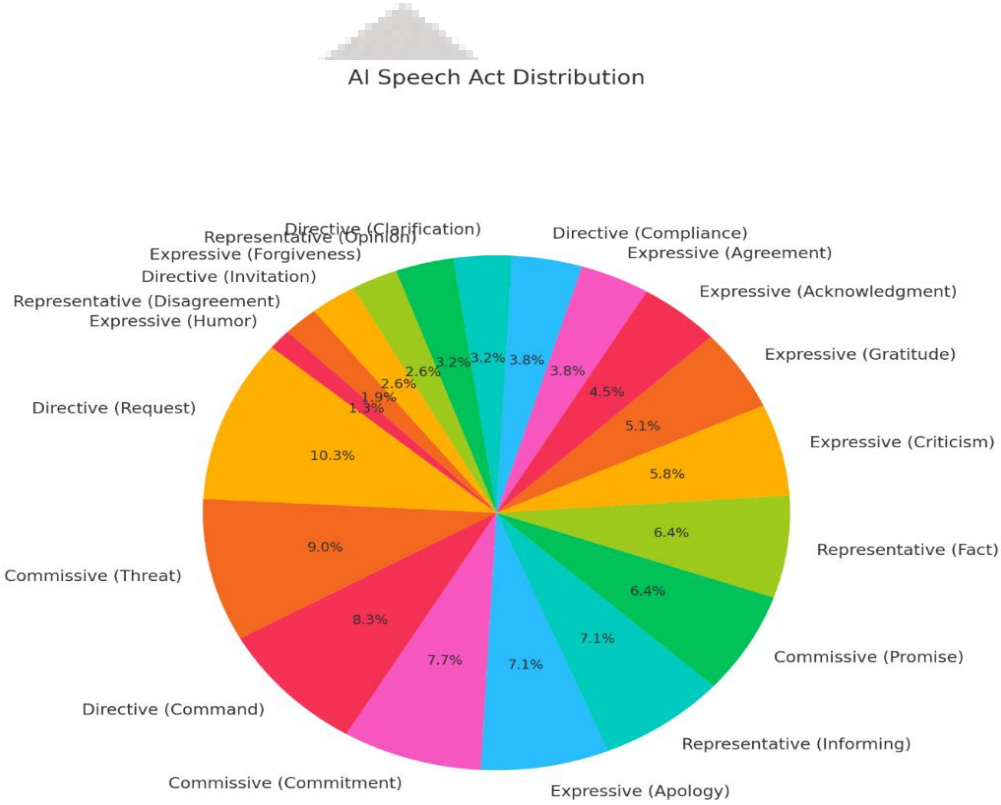

Figure 2 AI Speech Act Distribution

Expressive acts conveying psychological states or attitudes, like (Apology) is notably frequent at 7.1%, potentially reflecting a programmed politeness or error-handling strategy rather than genuine remorse. Others like Gratitude (5.1%) and Acknowledgment (4.5%), they are present but less dominant. They are unlike human interaction patterns. Concerning Criticism (5.8%), it appears more frequently. They may be linked to evaluative functions in specific AI applications.

Concerning conveying of information, representation is (7.1%) and Fact (6.4%). While information provision is a core AI function. These categories are less than the directive and commissive acts. Opinion (3.2%) and Disagreement (1.9%) are the lower in the spectrum . This might indicate a designed tendency for AI to avoid subjective stances or overt conflict.

Expressive categories like Humor (1.3%) and Forgiveness (2.6%) are among the least common. The AI's current limitations in generating sophisticated emotional and relational communication are highlighted . Similarly, Directive (Invitation) at 2.6% is relatively rare. Summarily , the AI speech act distribution portrayed in the chart refer to communication model that are optimized for function, procedure, and control. The prevalence of directives and commissive, suggests that AI dialogue potentially lacks the following the contextual sensitivity, the relational depth, and the pragmatic flexibility that human beings owned. This distribution highlights the ongoing challenge of imbuing AI with not just fluency, but realcommunicative competence.

7.3. ComparingAIandHumanSpeechAct Distributions

This section presents a comparison between the speech act distributions observed in AI-generated conversations and human conversations, based on the two pie charts previously analyzed. A table summarizing the frequency differences is provided, followed by an essay discussing the key distinctions and their implications.

Table 4 Comparative Table: AI vs. Human Speech Act Frequencies

Speech Act Category	AI Frequency (%)	Human Frequency (%)	Key Difference
Directives			

Request	10.3	6.9	AI uses significantly more requests.
Command	8.3	5.1	AI uses commands more frequently.
Compliance	4.5	4.0	Similar frequency.
Clarification	3.2	4.0	Humans use slightly more clarifications.
Invitation	2.6	3.4	Humans use slightly more invitations.
Commissives			
Threat	9.0	5.7	AI uses threats much more often.
Commitment	7.7	8.6	Humans show slightly higher use of commitments.
Promise	6.4	4.6	AI uses promises more often.
Expressives			
Apology	7.1	8.0	Humans use apologies slightly more.
Criticism	5.8	4.6	AI uses criticism slightly more.

Gratitude	5.1	5.7	Humans use gratitude slightly more.
Acknowledgment	4.5	6.3	Humans use acknowledgments noticeably more.
Agreement	3.8	6.9	Humans use agreement significantly more.
Forgiveness	2.6	3.4	Humans use forgiveness slightly more.
Humor	1.3	2.9	Humans use humor more than twice as often as AI.
Representatives			
Informing	7.1	7.4	Similar frequency.
Fact	6.4	4.0	AI states facts more often.
Opinion	3.2	5.1	Humans express opinions more often.
Disagreement	1.9	3.4	Humans express disagreementmore often.

Holding a kind of a comparative analysis of the speech act distributions in AI-generated and human conversations reveals fundamental differences in communication styles and priorities. Both of AI and humans utilize a range of speech acts. The frequency and balance of these acts deviate significantly.it highlights the

existent gap between artificial linguistic output and human interaction.The contrast lies in the occurrence of assertive and procedural acts in AI dialogue. It is unlike human where one can find relationality and balance. AI demonstrates a strong reliance on Directives, particularly Requests (10.3% vs. 6.9% in humans) and Commands (8.3% vs. 5.1%). Furthermore, AI employs Commissive acts like Threats (9.0% vs. 5.7%) and Promises (6.4% vs. 4.6%) more frequently than humans. This suggests that an AI communication model geared towards task completion. It simulates agency and prioritizes function over social nuance. Human communication, as shown in the chart, it emphasizes relational aspects. Humans utilize Expressive acts associated with social bonding and harmony more often, such as Agreement (6.9% vs. 3.8% in AI) and Acknowledgment (6.3% vs. 4.5%). Apologies are relatively frequent in both (8.0% human vs. 7.1% AI). The wider context of human Expressives represents a greater focus on interpersonal dynamics. Generally, humans express Opinions (5.1% vs. 3.2%) and Disagreements (3.4% vs. 1.9%) more readily, indicating a higher degree of subjective expression and engagement.

As for Humar,humans pay more attention on humor twice than that of AI (2.9% vs. 1.3%), AI faces context-dependent form of social expression. Both AI and humans frequently use Representative acts like Informing (7.1% AI vs. 7.4% human), AI tends to state Facts more often (6.4% vs. 4.0%), whereas humans lean more towards expressing Opinions. Essentially, the comparison held suggests that AI communication is currently characterized by a more directive, assertive, and sometimes formulaic approach, likely stemming from its design objectives focused on task execution and information delivery. In contrast, Human communication demonstrates a richer, more balanced pragmatic profile, i.e. integrating informational goals with reference to the social interaction, emotional expression, and relationship management.

8. Findings

The analysis done showed that there is major differences in speech act usage, pragmatic appropriateness, and adherence to conversational norms. The differences included four

interrelated dimensions: speech act distribution, functional deployment, maxim violation frequency, and overall communicative intent. First, there is a significant difference between AI and human speakers. AI responses revealed a strong preference for Directive and Commissive acts, specifically, requests, commands, threats, and promises. Moreover, these categories reflected the procedural and task-oriented nature of most chatbot systems. This indicates that they are often designed to elicit user action or simulate agency. Contrastively , human showed a more balanced distribution across speech act types, with the use of Expressive (e.g., apologies, gratitude) and Representative (e.g., informing, opinions) acts. This suggests that human dialogue is more contextually adaptive and socially situated, incorporating both affective and epistemic functions.

Second, AI chatbots have frequently defaulted to formulaic or scripted politeness strategies, especially through apologies and clarifications, often lacking contextual sensitivity. Moreover, their commissive acts are have been used without genuine intentional grounding, raising questions about the legitimacy of promises or threats when issued by non-sentient agents. For human responses, they have demonstrated intentional variability, emotional attunement, and a measured use of directives, often reflecting underlying social norms such as politeness, reciprocity, or face-saving.

Third, There has been shown a high maxim violation rates in AI-generated discourse. Over 85% of AI responses breached one or more of Grice's conversational maxims, most commonly the Maxims of Quality (false or unverifiable claims), Manner (ambiguity), and Quantity (over- or under-information). This means that these violations result from the AI's lack of real-world grounding. In other words, the violations can be attributed to its reliance on probabilistic language generation, and the absence of shared conversational context. In contrast, human responses have shown near-perfect maxim adherence, demonstrating a high degree of pragmatic control, relevance, and truthfulness.

Fourth, AI tends to struggle particularly with speech acts such as requests, threats, and apologies—areas where human speakers generally perform with minimal violations by relying on nuanced intent and real-time social reasoning. So, what's the bottom line? AI tends to play it safe in conversations, focusing on being clear and avoiding mistakes. This often makes it sound a bit stiff, overly careful, or even bossy. People, on the other hand,

naturally juggle getting information across, handling emotions, and being sensitive to the person they're talking to.

Conclusion

Basically, this research shows that while AI chatbots can string words together correctly and stay on topic, they really miss the mark when it comes to the feel of a real conversation – the pragmatic side of things. They often break the unspoken rules of chat (like Grice's maxims) and use fancy-sounding phrases that just don't fit the situation. There's a real gap between how AI talks and what people expect from a conversation.

AI talks the way it does because it's built for efficiency and avoiding risk. But that focus means it loses out on understanding emotions, being flexible in different situations, and building any real connection. When AI tries to give commands, make threats, or apologize, it often feels hollow because there's no real feeling or social understanding behind it. It highlights how tricky it is to teach machines the ethical and social nuances of human chat.

People in the study, however, showed they naturally use conversation strategically. They're clear when needed, hold back when appropriate, and follow social norms. How humans talk shows an awareness of the real world, emotional intelligence, and an ability to adapt to who they're talking to. Ultimately, even though AI can produce text that makes sense grammatically and topically, it often fails the 'common sense' test of conversation – breaking basic rules about being truthful, clear, relevant, and appropriate. Real conversational smarts aren't just about generating words; they're about understanding intentions and the situation as it unfolds.

Recommendations

To help AI get better at the pragmatic side of conversation, here are a couple of key things researchers should focus on:

1. researchers have to concentrate on boosting Context Smarts: Help AI get better at understanding what's happening right now in a conversation, not just relying on past data.
2. researchers have to focus on improving how AI handles things like apologies and using softer language (hedging) appropriately.

References

Austin, J. L. (1962). *How to do things with words*. Oxford University Press.

Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5185–5198. https://doi.org/10.18653/v1/2020.acl-main.463

Bojic, L., Kovacevic, P., & Cabarkapa, M. (2023). GPT-4 Surpassing Human Performance in Linguistic Pragmatics. *arXiv preprint* arXiv:2312.09545. https://arxiv.org/abs/2312.09545

Chaves, A. P., & Gerosa, M. A. (2021). How should my chatbot interact? A survey on human-chatbot interaction design. *International Journal of Human-Computer Interaction*, 37(10), 906–931. https://doi.org/10.1080/10447318.2020.1841438

Eragamreddy, N. (2025). The impact of AI on pragmatic competence. *The Journal of Teaching English for Specific and Academic Purposes*, 13(1), 169–189.

Georgila, K., Artstein, R., & Traum, D. (2020). Pragmatic appropriateness of chatbot responses in task-based dialogue. *Dialogue & Discourse*, 11(1), 55–83.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 20(11), 818–829.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41–58). Academic Press.

Henderson, P., Hu, J., Romoff, J., Brunskill, E., Jurafsky, D., & Pineau, J. (2018). Ethical challenges in data-driven dialogue systems. *Proceedings of AAAI/ACM AI Ethics Workshop*. https://aaai.org/ocs/index.php/WS/AAAIW18/paper/view/17802

Huang, J., He, H., & Zhu, M. (2019). Challenges in evaluating conversational agents: The role of pragmatics. *Proceedings of NeurIPS Workshop on Conversational AI*.

Jurafsky, D., & Martin, J. H. (2021). *Speech and language processing* (3rd ed.) [Draft version]. https://web.stanford.edu/~jurafsky/slp3/

Kiela, D., Bulat, L., Lample, G., & Cho, K. (2019). Supervised multimodal bitransformers for classifying images and text. *Proceedings of EMNLP*, 413–423.

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: state of the art, current trends and challenges. *Multimedia Tools and Applications*, 82, 3713–3744.

Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.

Levinson, S. C. (2023). *Pragmatics in the age of artificial intelligence*. Cambridge University Press.

Makowski, S., & Levin, E. (2024). Beyond pattern matching: The limits of pragmatic understanding in large language models. *Proceedings of the Association for Computational Linguistics*, 42, 82–96.

Miehling, E., Theodorou, A., Weller, A., & Singh, P. (2024). Conversational Maxims for Human-AI Interactions. *Findings of the Association for Computational Linguistics: EMNLP 2024*.

Najem, A. S., & Abbas, N. F. (2025). A pragmatic analysis of emotional blackmail in the American movie *No One Would Tell* (2018). *Waset Journal of Human Sciences*, 21(1/Pt1), 983–951.

Panfili, L., Duman, S., Nave, A., Phelps Ridgeway, K., Eversole, N., & Sarikaya, R. (2021). Human-AI interactions through a Gricean lens. *Proceedings of the Linguistic Society of America*, 6(1), 288–302.

Roller, S., Dinan, E., Goyal, N., Ju, D., Williamson, M., Xu, J., ... & Weston, J. (2021). Recipes for building an open-domain chatbot. *arXiv*, arXiv:2004.13637.

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

Searle, J. R. (1979). Expression and meaning: Studies in the theory of speech acts.Cambridge University Press.

Shieber, S. (2021). Pragmatic failures in dialogue systems: A case for human evaluation. *Journal of Artificial Intelligence Research*, 70, 623–648.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.

Wang, Z., Li, J., Ren, P., & Chen, Z. (2025). Cultural Learning-Based Culture Adaptation of Language Models. *arXiv preprint* arXiv:2504.02953.