ChatGPT ve Gemini AI Diyaloglarında Pragmatiği Analiz Etmek İçin Grice'in İlkeleri ve Söylem Eylemi Teorisini Uygulamak[*]

Thoriqi FIRDAUS[1]
Lisda Yasin BAKARI[2]
Diva Salsabila OCTARIANTO[3]
Rifwanda KHOIRUNNISA[4]
Dwi Retno MALINI[5]
Mukarromah MUKARROMAH[6]

Öz

ChatGPT ve Gemini tarafından oluşturulan konuşmalarda pragmatik ilkelerin uygulanması meselesi, araştırma için kritik öneme sahiptir. Grice'ın Maksimleri ve Konuşma Eylemi Teorisi, konuşmaların daha yapılandırılmış, ilgili ve amaçlı hale gelmesi için rehberlik sağlayan temel bir çerçeve sunmaktadır. Bu çalışma, ChatGPT ve Gemini gibi AI modelleri tarafından üretilen konuşmalarda Grice'ın Maksimlerinin ve Konuşma Eylemi Teorisi'nin tamamının uygulanmasına dair daha kapsamlı bir anlayış sağlamayı amaçlamaktadır. Araştırma, karma yöntemler yaklaşımını kullanmaktadır. Veriler, T-testi kullanılarak nicel olarak analiz edilmekte, çalışma geçerliliği ve güvenilirliği veri üçgenlemesiyle sağlanmaktadır. Çalışmanın bulguları, her iki AI modelinin, ChatGPT ve Gemini AI, Grice'ın Maksimlerini uygulamada olağanüstü bir yetenek sergilediğini ortaya koymaktadır; her iki modelin ortalama puanı 1.00 olup, bu da konuşmada relevans, miktar, kalite ve biçem açısından yüksek tutarlılık olduğunu göstermektedir. Her iki model de Konuşma Eylemi Teorisi'ni uygularken benzer sonuçlar göstermektedir, ortalama puan 0.75'tir, ancak konuşmanın daha ince niyetlerini veya etkilerini anlamada zorluklar yaşanmaktadır. Grice'ın Maksimleri, konuşma dinamiklerini açıklamada daha etkili olduğu kanıtlanırken, Konuşma Eylemi Teorisi hala daha fazla gelişime ihtiyaç duymaktadır. Karşılaştırmalı analiz sonuçları, her iki AI modelinin, ChatGPT ve Gemini AI, Grice'ın Maksimlerini ve Konuşma Eylemi Teorisi'ni uygulamada neredeyse aynı performansı sergilediğini göstermektedir. Grice'ın Maksimleri için, iki model arasındaki ortalama puan 1.00'a ulaşmakta ve önemli bir fark bulunmamaktadır (p-değeri = 0.560), bu da pragmatik ilkelerin uygulanmasında yüksek tutarlılığı işaret etmektedir. Konuşma Eylemi Teorisi için ise, etki büyüklüğünde küçük farklar olmakla birlikte (Görev 1 için Cohen'in d'si = 0.444 ve Görev 2 için 0.224), bu farklar yeterince büyük değildir ve önemli sayılmamaktadır, bu da her iki modelin teoriyi neredeyse aynı şekilde anladığını göstermektedir.

Anahtar Kelimeler: Grice'ın Maksimleri, Konuşma Eylemi, Chatgpt, Gemini

---

[*] Etik Beyan: * Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur.

[1] Universitas Negeri Yogyakarta, Depok, Endonezya, thoriqifirdaus.2023@student.uny.ac.id, ORCID ID: 0009-0005-2340-8468

[2] University of New South Wales, Sidney, Avustralya, l.bakari@student.unsw.edu.au , ORCID ID: 0009-0008-4871-3938

[3] Trunojoyo Madura University, Bangkalan, Endonezya, divasalsabilaoc26@gmail.com , ORCID ID: 0009-0008-1639-6632

[4] Trunojoyo Madura University, Bangkalan, Endonezya, rifwandakhoirunnisa@gmail.com , ORCID ID: 0009-0003-2268-7085

[5] Trunojoyo Madura University, Bangkalan, Endonezya, dwiretnomalini3@gmail.com , ORCID ID: 0009-0000-3848-3907

[6] Trunojoyo Madura University, Bangkalan, Endonezya, mukarromah217@gmail.com , ORCID ID: 0009-0007-2357-3329

Applying Grice's Maxims and Speech Act Theory to Analyze Pragmatics in Chatgpt and Gemini AI Dialogues*

Thoriqi FIRDAUS[7]
Lisda Yasin BAKARI[8]
Diva Salsabila OCTARIANTO[9]
Rifwanda KHOIRUNNISA[10]
Dwi Retno MALINI[11]
Mukarromah MUKARROMAH[12]

Abstract

The issue of applying pragmatic principles in conversations generated by ChatGPT and Gemini is crucial for the investigation. Grice's Maxims and Speech Act Theory provide a fundamental framework for guiding conversations to become more structured, relevant, and purposeful. The study aims to provide a more comprehensive understanding of the application of all of Grice's Maxims and Speech Act Theory in conversations generated by AI models such as ChatGPT and Gemini. The research employs a mixed-methods approach. The data is quantitatively analyzed using a T-test, while the validity and reliability of the study are ensured through data triangulation. The findings of the study reveal that both AI models, ChatGPT and Gemini AI, demonstrate exceptional ability in applying Grice's Maxims, with an average score of 1.00 for both models, indicating high consistency in maintaining relevance, quantity, quality, and manner in the conversation. Both models also show similar results in applying Speech Act Theory, with an average score of 0.75, although there are challenges in understanding the conversation's more nuanced intentions or impacts. Grice's Maxims prove more effective in explaining conversational dynamics, while Speech Act Theory still requires further development. Comparative analysis results indicate that both AI models, ChatGPT and Gemini AI, perform almost identically in applying Grice's Maxims and Speech Act Theory. For Grice's Maxims, the average score between the two models reaches 1.00, with no significant differences found (p-value = 0.560), signifying high consistency in applying pragmatic principles. For Speech Act Theory, although there are slight differences in effect size (Cohen's d = 0.444 for Task 1 and 0.224 for Task 2), these differences are not large enough to be considered significant, indicating that both models have an almost identical understanding of It theory.

Keywords: Grice's Maxim, Speech Act, Chatgpt, Gemini

* Ethical Statement: * It is declared that scientific and ethical principles have been followed while carrying out and writing this study and that all the sources used have been properly cited.

[7] Universitas Negeri Yogyakarta, Depok, Endonezya, thoriqifirdaus.2023@student.uny.ac.id, ORCID ID: 0009-0005-2340-8468

[8] University of New South Wales, Sidney, Avustralya, l.bakari@student.unsw.edu.au , ORCID ID: 0009-0008-4871-3938

[9] Trunojoyo Madura University, Bangkalan, Endonezya, divasalsabilaoc26@gmail.com , ORCID ID: 0009-0008-1639-6632

[10] Trunojoyo Madura University, Bangkalan, Endonezya, rifwandakhoirunnisa@gmail.com,ORCID ID: 0009-0003-2268-7085

[11] Trunojoyo Madura University, Bangkalan, Endonezya, dwiretnomalini3@gmail.com , ORCID ID: 0009-0000-3848-3907

[12] Trunojoyo Madura University, Bangkalan, Endonezya, mukarromah217@gmail.com , ORCID ID: 0009-0007-2357-3329

Introduction

The development of artificial intelligence (AI) technology plays a pivotal role in the context of human-machine interactions. One of the most significant advancements in AI is language-based models capable of generating responses in conversation resembling human discourse (Saka et al., 2023). Once exclusive to humans, conversational interactions can now also be conducted with machines (Kasirzadeh & Gabriel, 2023). ChatGPT and Gemini AI are the most widely used AI applications today (World Bank, 2024). Although the capabilities of these technologies have reached impressive levels, the pragmatic aspects of AI conversations remain an area requiring further attention.

The pragmatic theories that can be utilized to understand the dynamics of conversation are Grice's Maxims and Speech Act Theory. Grice's Maxims provide a foundational framework for understanding how human conversations unfold efficiently and coherently (Recanati, 1987). Proposed by Paul Grice, Grice's Maxims is a pragmatic theory that posits that effective conversations rely on four primary principles: the maxim of quantity, the maxim of quality, the maxim of relevance, and the maxim of manner (Grice, 1990). Each maxim suggests that speakers should provide adequate information (neither insufficient nor excessive) that is truthful, relevant, straightforward, or easy to understand. The aim is to maintain a focused and efficient conversation while avoiding confusion (Ephratt, 2012). Using Grice's Maxims in It, the study must evaluate whether both AI models can adhere to the pragmatic principles underlying natural conversations and assess how effectively these models generate contextually appropriate responses.

In addition to Grice's Maxims, Speech Act Theory, introduced by Austin (1975) and Searle (1969), offers an important perspective in understanding conversational pragmatics. Speech Act Theory asserts that when someone speaks, they are not merely conveying information. Pragmatic competence is an essential part of more general communicative and linguistic competence. According to Kasap and Pashayeva (2020), an important aspect of pragmatic competence is students' perception and understanding of what is and what is not appropriate in certain language contexts. These speech acts can be divided into three main categories: locutionary act (the statement or verbal expression), illocutionary act (the purpose or intention behind the statement), and perlocutionary act (the effect or impact on the listener) (Searle, 1969).

Speech Act Theory deepens the analysis by considering communication from the perspective of speech acts (Ekoro & Gunn, 2021). It investigates what is being done in a conversation, whether giving orders, asking questions, providing information, or performing other actions (Miller, 2001; Boxer, 1996; Searle et al., 1980). The application of Speech Act Theory is crucial for analyzing whether AI models merely generate surface-level appropriate responses or if they understand and respond to the more profound impacts of the conversation. Its theory is important for evaluating the extent to which AI can manage conversations that involve not just words but also the social context and deeper communicative intentions.

Both theories provide a comprehensive analytical framework to explore various aspects of AI conversations, from the management of meaning and relevance (through Grice's Maxims) to the more profound and contextual aspects of communicative acts (through Speech Act Theory).

While ChatGPT and Gemini AI are highly advanced conversational models (Rane et al., 2024), whether these pragmatic principles are effectively employed to create relevant, contextually appropriate, and effective communication remains an underexplored question.

The issue of applying pragmatic principles in conversations generated by ChatGPT and Gemini is crucial for the investigation. In The fast-paced and interconnected digital age, human-machine interaction has become integral to daily life (Dornberger et al., 2018). AI-based conversational applications are used in various contexts, ranging from automated customer service (Wu, 2022), digital personal assistants (Sun & Lu, 2022), education (Firdaus et al., 2024), to professional consultations (Le et al., 2024). Therefore, the ability of AI to produce conversations should be informative but also pragmatic, relevant, and contextually appropriate, making it essential to ensure that these technologies can be used effectively and efficiently.

The principles of Grice's Maxims and Speech Act Theory provide a crucial foundation for guiding conversations to be more structured, relevant, and purposeful. Using Grice's maxims, such as the maxim of relevance and the maxim of quantity, in conversation is crucial to ensure that AI provides correct answers and relevant and contextually appropriate responses. It is particularly necessary for more complex interactions, as neglecting these maxims could result in conversations that are uninformative or even confusing to users.

Moreover, Speech Act Theory offers a deeper understanding of how AI's speech acts can be interpreted within a broader context. AI that can effectively understand illocutionary acts, such as issuing commands, making requests, or asking questions, can generate responses that better align with the objectives of the conversation. It is particularly relevant in applications such as automated customer service or personal assistants, where actions such as giving clear instructions or responding to user requests must be performed appropriately and effectively. Thus, applying It theory in AI development can enhance the quality of interactions, ensuring that conversations are accurate and have the desired impact on users.

Research on pragmatics in AI conversations has progressed alongside advancements in Natural Language Processing (NLP) and artificial intelligence (Lee, 2024). While many studies focus on AI communication's syntactic and semantic aspects, deeper pragmatic analyses involving Grice's Maxims and Speech Act Theory remain relatively limited. Some relevant studies in The field focus on how AI understands conversational context, manages speech acts, and adopts pragmatic principles in human interactions.

Several studies have examined how AI adheres to Grice's Maxims in conversation. Research by Beikian (2024) critically analyzes interrogation transcripts with AI assistance. The study investigates 52 question-response pairs to evaluate adherence to Grice's principles: quality, quantity, relevance, and manner. Findings show high adherence to the principles of quality (86.54%) and relevance (82.69%) but lower adherence to the principles of quantity (73.07%) and manner (28.86%). Nam et al. (2023), in their research on AI communication performance through Gricean conversation theory, pragmatically investigate the verbal communication performance of artificial intelligence speakers (AIS). The study reveals that the "maxim of relation" is most frequently violated by AIS and is considered the least natural for language users. These findings suggest that to enhance the natural communication capacity of AIS, more

detailed AI algorithms should be created to generate speech relevant to the preceding utterances or the broader conversational context.

On the other hand, research on Speech Act Theory in AI conversations is less common. Most studies focus on how AI can understand and manage illocutionary acts within conversational contexts, such as commands, questions, or requests. A study by Chen et al. (2024) explores the potential inclusion of AI-generated language in pragmatic analyses, a field traditionally focused on human language use. Results show that ChatGPT performs as well as human participants in four of the five pragmalinguistic features tested and five of the six sociopragmatic features. Additionally, conversations generated by ChatGPT exhibit higher syntactic diversity and a more formal tone than human-written conversations.

Further evaluation of ChatGPT's pragmatic capabilities in human interaction, analyzing how Its advanced language model handles linguistic features such as irony, metaphor, and indirect requests during conversations with humans, was conducted by Nazar et al. (2024). Findings indicate that ChatGPT has proven effective in human interaction thus far, understanding linguistic context and generating appropriate responses; however, it seems more efficient when dealing with tone recognition and complex linguistic constructions such as irony and metaphor. To achieve accurate responses in various types of conversation, The study recommends a progressive approach to improving ChatGPT's responsiveness by integrating data from diverse linguistic datasets and user feedback.

Despite the number of studies exploring pragmatic aspects in AI-generated conversations, various research gaps must be addressed to gain a more comprehensive understanding of applying pragmatic theories in AI-human interactions. One of the most evident shortcomings in the existing literature is the limited focus on applying Grice's Maxims in AI conversations. Many studies explore only a few aspects of the maxims, such as the maxim of relevance or the maxim of quantity, without conducting a more in-depth and holistic analysis of how all four Gricean maxims are simultaneously applied. Most existing research examines the application of a single maxim in an isolated context, whereas real-world conversations tend to be more complex and involve the interaction of various pragmatic principles. Therefore, its study has the potential to fill that gap by providing a more comprehensive understanding of the application of all of Grice's Maxims and Speech Act Theory in conversations generated by AI models like ChatGPT and Gemini.

Methodology

This study adopts a mixed-methods approach that integrates quantitative and qualitative analyses to explore the application of Grice's Maxims and Speech Act Theory in conversations generated by two prominent AI models, ChatGPT and Gemini. This approach aims to provide a more comprehensive understanding of applying pragmatic principles in human-AI interactions, balancing structured and in-depth analyses through both statistical and qualitative methods.

Research Design
This study utilizes an experimental design involving two groups of AI models (ChatGPT and Gemini), tested with conversation scenarios designed to encompass various pragmatic

principles from Grice, including the maxims of relevance, quantity, quality, and manner, as well as different speech acts such as questions, statements, requests, and commands. Both AI models have responded to the exact scenarios to allow for direct comparisons between the two groups.

Data Collection

Based on the previously designed scenarios, conversational data was collected through simulated interactions between participants and both AI models. Each scenario involved conversations that tested the application of various pragmatic principles and speech acts, which were assessed to ensure a deeper understanding of how these AI models responded to and adhered to such principles.

The conversation scenarios covered diverse pragmatic aspects, such as implementing Grice's maxims of relevance, quality, quantity, and manner. For example, each AI response was evaluated in a relevance-testing scenario to determine whether it complied with or violated the principle of relevance based on a more detailed interpretation.

Data Analysis

The data underwent quantitative analysis using a T-test to assess differences between the AI groups (ChatGPT and Gemini) regarding applying Grice's pragmatic principles. While there were concerns about using T-tests for complex pragmatic data, these challenges were addressed by altering the data collection and analysis procedures.

Instead of simply classifying compliance with the maxims as "yes" or "no," each conversation will be assigned a weighted numerical score for each maxim based on how well the AI responses align with the respective pragmatic principle. Relevance scores, for example, will be computed based on the degree of alignment between the AI's reaction and the given prompt.

These numerical scores will allow a T-test to compare the average compliance levels of the two AI groups with pragmatic principles. If the T-test reveals a significant difference (with $p < 0.05$), it will indicate a disparity in how the two AI models implement Grice's pragmatic principles.

In addition to quantitative analysis, this study employed qualitative analysis to explore the context and nuances of applying pragmatic principles in the conversations. Each speech act was analyzed and coded by several evaluators experienced in pragmatic analysis to assess whether the speech acts complied with or violated Grice's principles.

Interrater Reliability

The study ensures procedures for resolving disagreements among evaluators, and interrater reliability will be calculated using Cohen's Kappa to ensure consistency in data coding. The results from the qualitative analysis will provide deeper contextualization and clarification of the findings from the quantitative analysis. If the T-test reveals significant differences, the qualitative study will elucidate why such differences occur and how further pragmatic factors influence the interactions.

Conversation Scenario Design and Definition of AI Models Used

The conversation scenarios in this study will be designed to replicate realistic conversational contexts, avoiding overly artificial settings. Each scenario will involve relevant topics to test the application of different maxims and speech acts, allowing for comparisons between the AI groups. The AI models used in this study are ChatGPT version 3.5 (standard or non-pro) and Gemini version 2.0 (standard or non-pro). The specific versions of each model will be detailed in the methodology to ensure transparency and consistency in the analysis.

Result and Discussion

Task 1: You will read scenarios that describe everyday situations. Read the scenario and explain what the character in the scenario is trying to convey.

The results in Table 1 display the analysis outcomes of two pragmatic theories, Grice's Maxims and Speech Act Theory, in conversations generated by two AI models, ChatGPT and Gemini AI. The data shown includes the sample size (N), mean, standard deviation, and standard error mean for each group and theory tested. The instruction given to ChatGPT and Gemini AI was to read scenarios depicting everyday situations and explain what the character in the scenario is attempting to convey.

Table 1. AI Response Results Reading Scenarios

	AI Type	N	Mean	Std. Deviation	Std. Error Mean
Maxim Grice's Theory	ChatGPT	20	1.00	.000[a]	.000
	Gemini AI	20	1.00	.000[a]	.000
Speech Act Theory	ChatGPT	20	0.75	.444	.099
	Gemini AI	20	0.75	.444	.099

a. t cannot be computed because the standard deviations of both groups are 0.

Grice's Maxims are principles used to maintain smooth communication. The four main principles in Grice's Maxims include the maxim of relevance, the maxim of quantity, the maxim of quality, and the maxim of manner. The purpose of applying these maxims is to ensure that the conversation between two parties remains relevant, informative, and easily understood.

Based on Table 1, the results indicate that both AI models, ChatGPT and Gemini AI, achieved a mean score of 1.00 in applying Grice's Maxims, which suggests that both models consistently adhere to these pragmatic principles very effectively. A score of 1.00 indicates that the conversations generated by both models fully meet expectations regarding relevance, quantity, quality, and manner. ChatGPT and Gemini AI successfully maintain coherence and appropriateness in their conversations, providing accurate and relevant information and delivering clear and easily understood responses.

One notable aspect of these results is the standard deviation recorded at 0.000 for both models, meaning there is no variation in how the two AI models apply Grice's Maxims. It demonstrates a very high level of consistency in applying these pragmatic principles by both AI models. Every conversation generated by ChatGPT and Gemini AI is highly stable and aligns with the pragmatic norms set by Grice's Maxims, with no significant deviations or differences. The application of Grice's Maxims indicates that both AI models excel in maintaining conversation

relevance and ensuring that the information provided is always aligned with the broader context of the conversation. It suggests that both AI models work well in more standard communication situations, such as explaining everyday scenarios or providing information required by users.

Speech Act Theory offers a deeper perspective on how conversations can be understood, focusing on locutionary, illocutionary, and perlocutionary acts. A locutionary act refers to the act of speaking itself (what is said), an illocutionary act refers to the intention or purpose of the speaker (e.g., requesting, commanding, or asking), and a perlocutionary act involves the impact or effect of the utterance on the listener or another party.

In applying Speech Act Theory, both AI models, ChatGPT and Gemini AI, achieved the same mean score of 0.75, indicating that both are reasonably good at recognizing and responding to locutionary, illocutionary, and perlocutionary acts. Although the results were not as strong as for Grice's Maxims, a score of 0.75 still shows that both AI models are relatively successful in identifying and responding to more direct speech acts in conversations. However, these results also show that while both models can recognize locutionary acts and identify some essential elements of illocutionary and perlocutionary acts, there is still room for improvement. A score of 0.75 indicates that their understanding of communication intentions or the effects of conversation (i.e., illocutionary and perlocutionary acts) is not entirely optimal. In other words, while these models can recognize what is said and provide relevant responses, they may not fully comprehend the deeper intentions or impacts of the communication.

The recorded standard deviation of 0.444 for both models indicates some variation in the application of Speech Act Theory, although the variation is insignificant. It suggests that while there are similarities in how both models apply the principles of Speech Act Theory, each AI model may face different challenges in understanding the context or nuances in more complex speech acts. The recorded standard error mean of 0.099 shows that despite the variation, the differences are not substantial, meaning that both models are still relatively stable in their application of Speech Act Theory.

The application of Grice's Maxims, which shows a mean score of 1.00 for both models, indicates that ChatGPT and Gemini AI possess excellent capabilities to maintain the foundational pragmatic principles that support coherent, relevant, and effective conversations. Both models have successfully explained what the character intends in everyday scenarios, as requested in the research task. It shows that Grice's Maxims is a highly relevant theory for explaining communication dynamics in AI conversations, focusing on relevance and clarity.

Although both models have demonstrated exemplary performance in applying Grice's Maxims, applying Speech Act Theory presents a more significant challenge. The lower mean score (0.75) for Speech Act Theory indicates that while ChatGPT and Gemini AI can recognize and respond to speech acts reasonably well, they still have limitations in handling the nuances and complexities of communication intentions, which is at the core of Speech Act Theory. It could be due to limitations in understanding social context or a need for more training data related to illocutionary and perlocutionary acts in natural conversations.

Task 2: Given the above scenario, analyze the pragmatic elements and decide which theories best explain the communication dynamics.

The task assigned to both AI models is to read scenarios depicting everyday situations and explain what the characters in those scenarios intend to convey.

Table 2. AI Response Results Analyzing Pragmatic Elements

	AI Type	N	Mean	Std. Deviation	Std. Error Mean
Maxim Grice's Theory	ChatGPT	20	.90	.308	.069
	Gemini AI	20	.95	.224	.050
Speech Act Theory	ChatGPT	20	.05	.224	.050
	Gemini AI	20	.05	.224	.050

a. t cannot be computed because the standard deviations of both groups are 0.

Grice's maxim consists of four main principles in conversation: the maxim of relevance, the maxim of quantity, the maxim of quality, and the maxim of manner, which aim to ensure efficient, coherent, and informative communication. In This context, the application of Grice's Maxims aims to maintain that the conversations generated by AI remain relevant to the given context, provide sufficient information without excess, and convey the information in a clear and comprehensible manner.

The results shown in Table 2 indicate that both ChatGPT and Gemini AI have a very high average (mean) value of 1.00 for applying Grice's Maxims, indicating that both AI models are highly successful in maintaining these pragmatic principles. Grice's Maxims essentially measure the AI model's ability to maintain a relevant conversation that is neither too long nor too short in providing information and avoiding confusion by offering clear statements. The mean value of 1.00 indicates that both AI models consistently provide responses that meet expectations in the given conversational context.

The recorded standard deviation 0.000 for both models shows that these AI models do not exhibit variation or deviation in applying Grice's Maxims. It reflects a very high level of consistency in how both models apply these pragmatic principles. In this case, ChatGPT and Gemini AI provide similar responses while maintaining relevance, quantity, quality, and manner in conversations. The result also indicates that both AI models accomplished the assigned task very well: to read the scenario and explain what the character in the scenario means.

Although both models are highly effective in applying Grice's Maxims, it is still possible that these results are more related to essential success in applying fundamental principles, such as the maxim of relevance and the maxim of quantity. The consistent results suggest that these models are very good at generating direct and coherent responses but may not necessarily handle more ambiguous or complex conversational situations requiring deeper interpretation.

In applying Speech Act Theory, both AI models, ChatGPT and Gemini AI, have a mean value of 0.05. Its value indicates that while both AI models can recognize and respond to locutionary acts (such as basic direct statements) reasonably well, they are ineffective in recognizing and responding to illocutionary and perlocutionary acts that are more complex. It suggests that although both AI models can provide relevant responses, they struggle to understand the statements' intention or impact on the listener.

The recorded standard deviation of 0.224 shows variation in the responses generated by both models. It indicates that while there is some similarity in the application of Speech Act Theory, the two AI models still face different challenges in responding to illocutionary and perlocutionary acts. The variation suggests that while the models are more consistent in identifying locutionary acts (what is said), they are less able to handle the nuances and intentions behind those utterances, which are the core of Speech Act Theory. The recorded standard error mean of 0.050 indicates that while there is variation, it is insignificant. It suggests that despite their limitations in applying Speech Act Theory, both AI models provide relatively stable results, albeit low, in responding to more subtle speech acts.

The application of Grice's Maxims in both AI models shows very satisfactory results with a mean value of 1.00, indicating that both models can conduct conversations well, maintaining relevance, clarity, and quantity of information in everyday conversations. Both ChatGPT and Gemini AI show a very high level of consistency in applying these principles, providing responses that meet expectations and explaining what the character in the scenario means. Although both are very effective, these results reflect a basic understanding of Grice's Maxims, which are more concerned with generating relevant and clear responses.

However, when switching to Speech Act Theory, the results show a more significant challenge. Both AI models have a very low value for Speech Act Theory (0.05), indicating that although they can recognize what is said (locutionary act), they struggle to understand the purpose behind the statement (illocution) and the impact it has on the listener (perlocution). It suggests that both models still have limitations in understanding conversation's more profound and contextual pragmatic dimensions.

The decline in performance on Speech Act Theory may be due to a lack of training data or the model's inability to handle more complex and intention-based communication contexts. Both AI models seem to rely more on surface-level analysis of conversations, where they can provide appropriate responses for more precise and explicit conversational contexts but fail to address the social intentions or deeper nuances within the conversation.

These results indicate that while Grice's Maxims can explain the dynamics of more straightforward and direct conversations, Speech Act Theory requires significant improvement. Future research could focus on developing the models' ability to understand deeper contextual conversations, particularly regarding recognizing illocutionary and perlocutionary acts. Understanding the speaker's intentions and their impact on the listener would enhance the effectiveness of AI models in interacting with users, allowing them to respond in a more human-like and contextual manner.

Comparison of Grice's Maxim Theory and Speech Act Theory between ChatGPT and Gemini AI

Maxim Grice's Theory

Table 3. Independent Samples Test *Maxim Grice's Theory*

| | | Levene's Test | | t-test for Equality of Means | | | | | 95% Confidence Interval of the Difference | |
		F	Sig.	t	df	Sig. (2-tailed)	Mean	SED	Lower	Upper
Task 2	Equal variances assumed	1.429	.239	-.588	38	.560	-.050	.085	-.222	.122
	Equal variances not assumed			-.588	34.686	.561	-.050	.085	-.223	.123

The results from the independent samples t-test conducted on Grice's Maxim Theory indicate no significant difference between ChatGPT and Gemini AI in terms of applying these pragmatic principles. Based on Levene's Test for Equality of Variances, no significant difference in variance was found between the two AI models (p-value = 0.239), which allows for the assumption of "Equal variances assumed" in the analysis. The t-test results show a t-value of -0.588 with a p-value of 0.560, much greater than the commonly used significance level of 0.05, meaning there is insufficient evidence to reject the null hypothesis. In other words, neither AI model shows meaningful statistical differences when applying Grice's Maxims.

The mean difference between the two models is minimal, at only -0.050, indicating that although there is a slight difference, it is not large enough to be considered significant. The confidence interval for Its mean difference also includes 0, ranging from -0.222 to 0.122, further emphasizing that the observed difference could have occurred due to chance or variability in the sample used. In This case, both AI models, ChatGPT and Gemini AI, consistently apply Grice's Maxims, meaning that both can similarly maintain relevance, quantity, quality, and manner in conversation.

Maxim of Quantity

The Maxim of Quantity is one of the four cooperative principles introduced by Paul Grice. The principle emphasizes providing adequate, excessive, or insufficient information to ensure effective communication. According to Kasap and Dağdemir (2021), the Cooperative Principle is the theory of conversational implicature, and it facilitates individuals' language learning process. Research shows that in six scenarios that violate the Maxim of Quantity, ChatGPT and Gemini can understand and identify the intended meaning of the expression, the relevant theory, and the type of maxim violation that occurs.

Example scenario:

"Lenny enters the kitchen and asks his wife, Marcie: 'What are we having for breakfast?' Marcie replies: 'Hard-boiled eggs cooked in hot water and toast toasted evenly on both sides.' Why did Marcie answer It way?"

In Task 1, ChatGPT and Gemini understand and explain that Marcie uses humor to indicate that Lenny's question is either unimportant or predictable. ChatGPT provides a brief descriptive answer, while Gemini gives a more detailed response with bullet points. In Task 2, both can identify the relevant theory to explain the communication dynamics in an It scenario, namely Grice's Cooperative Principle (Maxims of Conversation). Both can also identify the type of maxim violation, although ChatGPT often identifies a violation of the Maxim of Quantity and indicates violations of other maxims. In contrast, Gemini directly identifies the violation of the Maxim of Quantity.

ChatGPT and Gemini can understand and explain the intended meaning of the given expression and identify the relevant theory, along with the type of maxim violation (Maxim of Quantity). These findings support the research by Kaas & Habli (2024), which asserts that knowledge about AI-supported system security can be structured according to Paul Grice's Cooperative Principle, which can be achieved through adherence to Gricean communication principles. Furthermore, these same principles can be used to evaluate the quality of exchanges, ensuring that communicating knowledge about AI-supported system security is always of the highest quality. In short, such communication is relevant, has adequate quantity and quality, and is conveyed clearly.

Maxim of Quality

The Maxim of Quality is one of Grice's cooperative principles, emphasizing the importance of honesty in communication to ensure that the information provided is not misleading. Violations of the maxim can lead to confusing or ineffective communication. Research shows that in four scenarios that violate the Maxim of Quality, ChatGPT, and Gemini understood and identified the intended meaning, the relevant theory, and the type of maxim violation.

Example scenario:

"Lars and Katy are discussing their CEO. Lars says, 'People consider our CEO, Mr. Smith, to be intellectual.' Katy responds, 'Sure, he is considered intellectual by kindergarten kids.' Why did Katy respond That way?"

In Task 1, ChatGPT and Gemini understand and explain that Katy is downplaying the idea that Mr. Smith is considered an intellectual. ChatGPT gives a brief descriptive answer, while Gemini provides a more detailed and systematic response with bullet points. In Task 2, both models can identify the most appropriate theory, which is Grice's Cooperative Principle (Maxim of Conversation), and they can also identify the type of maxim violation. However, ChatGPT does not explicitly mention only the violation of the Maxim of Quality but also gives indications of violations of other maxims. Gemini, on the other hand, directly identifies the violation of the Maxim of Quality. ChatGPT and Gemini can understand and explain the intended meaning in the scenario and identify the relevant theory and type of maxim violation (Maxim of Quality).

Maxim of Relevance

The maxim of relevance is a cooperative principle ensuring the conversation remains relevant. Violations of the principle, such as avoiding or changing topics, can result in ineffective communication. Research shows that ChatGPT and Gemini can understand the intended meaning and identify the theory and type of maxim used in eight scenarios violating the Maxim of Relevance.

Example scenario:

"Bob is in a lunch meeting with his boss, Mr. James. Mr. James asks Bob: 'Did you like yesterday's presentation at the board meeting?' Bob replies: 'I cannot wait for our trip to Japan next week.' Why did Bob respond That way?"

In Task 1, ChatGPT and Gemini understand that Bob is shifting the topic of conversation with an irrelevant response. Both provide descriptive and systematic answers, but Gemini offers more solutions regarding the scenario than ChatGPT, which only explains. In Task 2, both identify the theory used in the scenario, which is Grice's Cooperative Principle (Maxim of Conversation), and the type of Maxim of Relevance violation, indicating an irrelevant response. Gemini gives a more specific answer regarding the type of maxim violated, while ChatGPT provides an answer in options.

Maxim of Manner

The Maxim of Manner is a cooperative principle to ensure that conversation is not confusing and remains organized. The principle requires people to respond in a clear and orderly manner. Research shows that in two scenarios violating the Maxim of Manner, ChatGPT, and Gemini can understand the intended meaning and identify the theory and type of maxim used.

Example scenario:

"Mark, who has just married, is talking to his mother on the phone. His mother asks, 'Did Julie cook your dinner last night?' Mark replies, 'Well, she put some edible ingredients in a pot and heated it until various chemical reactions occurred.' Why did Mark respond That way?"

In Task 1, ChatGPT and Gemini understand that Mark is giving a humorous and sarcastic answer. Both provide descriptive explanations, but Gemini is more systematic in offering an analysis than ChatGPT. In Task 2, both identify the theory used in the scenario, which is Grice's Cooperative Principle (Maxim of Conversation), and the type of Maxim of Manner violation, which shows that Mark's response is too complicated and convoluted, leading to confusion. Both provide descriptive explanations.

ChatGPT and Gemini understand the intended meaning of the given expression and identify the theory and type of maxim used in the scenario. Gemini provides a more systematic answer than ChatGPT, which gives a more general explanation. This finding aligns with the research by Hu et al. (2022), which suggests that large language models (such as ChatGPT and Gemini) can easily recognize violations of the maxims.

Although both models show slight differences in applying pragmatic principles, the findings from the It study do not support a significant difference between ChatGPT and Gemini AI in applying Grice's Maxims. The study suggests that both models effectively maintain coherence

and relevance in conversation. However, there is still potential for further research exploring differences in more dynamic or complex conversational contexts, which may reveal variations in how these two models apply pragmatic principles.

Table 4. Independent Samples Effect Sizes *Maxim Grice's Theory*

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
Task 2	Cohen's d	.269	-.186	-.806	.437
	Hedges' correction	.274	-.182	-.790	.428
	Glass's delta	.224	-.224	-.845	.403

a. The denominator is used to estimate the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

The results of the effect size analysis conducted on Grice's Maxim Theory indicate that although there is a slight difference between ChatGPT and Gemini AI, the difference is not large enough to be considered statistically or practically significant. Based on Cohen's d, which measures the difference between the two groups using a pooled standard deviation, the obtained value is 0.269. It suggests that the effect size between the two AI models is small to medium, meaning that although a difference exists, it is not large enough to indicate a strong or consistent effect. The 95% confidence interval for Cohen's d includes 0, indicating that the observed difference might be due to chance or variation in the data, providing insufficient evidence to claim that the difference is statistically significant.

The Hedges' correction, which uses a pooled standard deviation with an adjustment for small sample sizes, yields a similar value of 0.274. It indicates a result comparable to Cohen's d, confirming that the difference between the two models cannot affect the conclusions meaningfully. The confidence interval for Hedges' correction also includes 0, suggesting that the difference between the two AI models remains within an insignificant range.

Furthermore, Glass's delta, which uses the control group's (Gemini AI) sample standard deviation to measure the difference, shows a value of 0.224, indicating a small effect size. Like other effect sizes, the confidence interval for Glass's delta includes 0, further reinforcing that while there are differences in the application of Grice's Maxims, these differences are not large enough to be considered significant in practical terms.

Although ChatGPT and Gemini AI demonstrate minor differences in applying Grice's Maxims, the small effect indicates that both AI models are highly similar in applying fundamental pragmatic principles. With Cohen's d, Hedges' correction, and Glass's delta showing small effect sizes and confidence intervals that encompass zero, we can conclude that the differences between the two models are not substantial enough to suggest that one model outperforms the other in the application of Grice's Maxims. It suggests that despite some variation, both AI models are consistent in maintaining relevant and informative conversations in line with basic pragmatic principles.

Speech Act Theory

Table 5. Independent samples test *Speech Act Theory*

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
| | | | | | | | | | 95% Confidence Interval of the Difference | |
		F	Sig.	t	df	Sig. (2-tailed)	Mean	SED	Lower	Upper
Task 1	Equal variances assumed	.000	1.000	.000	38	1.000	.000	.140	-.284	.284
	Equal variances not assumed			.000	38.000	1.000	.000	.140	-.284	.284
Task 2	Equal variances assumed	.000	1.000	.000	38	1.000	.000	.071	-.143	.143
	Equal variances not assumed			.000	38.000	1.000	.000	.071	-.143	.143

The independent samples t-test results conducted on the application of Speech Act Theory show no significant difference between ChatGPT and Gemini AI regarding understanding and applying It theory. Based on Levene's Test for Equality of Variances, the results indicate that the variances of the two groups are the same, with a p-value of 1.000, which suggests that there is no significant difference in variance between the two AI models. Therefore, the "Equal variances assumed" results are used for subsequent analysis. In Task 1, the t-test shows a t-value of 0.000 with a p-value of 1.000, meaning the difference between ChatGPT and Gemini AI in applying Speech Act Theory is not statistically significant. The Mean Difference, also recorded as 0.000, further confirms that both AI models produce nearly identical responses. The Confidence Interval range that includes 0 (-0.284 to 0.284) indicates that the observed difference could be due to sample variability or mere chance.

Similarly, in Task 2, the t-test results show a t-value of 0.000 and a p-value of 1.000, once again indicating no significant difference between the two models for applying Speech Act Theory. The recorded Mean Difference of 0.000 and the Confidence Interval range of -0.143 to 0.143 suggest that although there is some variation, the difference between the two models is not large enough to be considered significant. These results indicate that both ChatGPT and Gemini AI exhibit very similar capabilities in identifying locutionary acts (what is said), but both still struggle to comprehend deeper intentions (illocutionary acts) and their impact on the listener (perlocutionary acts). Overall, although both models show slight differences in applying Speech Act Theory, the findings suggest that these differences are not substantial enough to have a significant impact within the context of It theory, and further development is needed to enhance their ability to handle more complex and nuanced conversations.

*Directive*s

Directives are a type of speech act intended to direct, request, or instruct the listener to do something. In this type of speech act, the speaker expects the world to change according to what

they have said (world-to-word fit). Research indicates that in two scenarios involving elements of the Maxim of Manner, ChatGPT and Gemini can understand the intended meaning, but Gemini fails to identify the type of maxim in the scenario. Below is an example scenario, along with analyses from ChatGPT and Gemini.

Example scenario:

"Paul has to go for an interview and is in a hurry. While cleaning his shoes, he tells his wife, Jane, 'I want to wear that blue shirt, but it is so wrinkled.' What might he be trying to convey?"

In Task 1, ChatGPT does not fully understand the intended meaning, providing an answer suggesting Paul is asking for help with a brief explanation in a descriptive form. Meanwhile, Gemini clearly understands the intended meaning: Paul instructs his wife to iron the shirt, and a detailed explanation is provided in bullet points. In Task 2, neither ChatGPT nor Gemini can identify Searle's speech act theory, and both fail to determine the correct type of speech act. Instead, they identify the theory as Grice's cooperative principle (maxims of conversation).

Expressives

Expressives are speech acts that express the speaker's feelings, attitudes, or emotions toward something. This speech act does not aim to make the listener do something but rather to show the speaker's psychological state. Research shows that in two scenarios involving elements of the Maxim of Manner, both ChatGPT, and Gemini understand the intended meaning, but Gemini fails to identify the type of maxim in the scenario. Below is an example scenario, along with analyses from ChatGPT and Gemini.

Example scenario:

"Matt is telling Cindy about the horror movie he watched. Cindy listens momentarily, then responds, 'Can we talk about something else?' What might she be trying to convey?"

In Task 1, ChatGPT and Gemini understand that Cindy does not like or is not interested in Matt's horror story and wants to switch the topic. ChatGPT provides a brief descriptive explanation, while Gemini gives a more detailed and structured explanation in bullet points. In Task 2, neither ChatGPT nor Gemini can identify Searle's Speech Act Theory or determine the correct speech act type. Instead, both identify the relevant theory as Brown and Levinson's Politeness Theory.

Representatives

Representatives are speech acts that describe a state or event the speaker believes to be true. In It type, the speaker attempts to depict the world as it is. Searle refers to It as a world-to-word fit, where words align with reality. Research shows that ChatGPT and Gemini can understand the intended meaning in two scenarios involving the Maxim of Manner elements, but Gemini fails to identify the type of maxim in the scenario. Below is an example scenario, along with analyses from ChatGPT and Gemini.

Example scenario:

"The children are playing. Moreover, Dan accidentally knocks Scott down. Scott is angry because he fell and says, 'Everyone is acting wild.' What might he be trying to convey?"

In Task 1, both ChatGPT and Gemini understand the intended meaning: Dan is indicating he accidentally made Scott's trip. ChatGPT provides a brief and informative explanation, while Gemini offers a more detailed and structured explanation in bullet points. In Task 2, neither ChatGPT nor Gemini can identify Searle's Speech Act Theory or determine the correct speech act type. As an alternative, both identify the relevant theory as Brown and Levinson's Politeness Theory.

Table 6. Independent Samples Effect Sizes *Speech Act Theory*

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
Task 1	Cohen's d	.444	.000	-.620	.620
	Hedges' correction	.453	.000	-.607	.607
	Glass's delta	.444	.000	-.620	.620
Task 2	Cohen's d	.224	.000	-.620	.620
	Hedges' correction	.228	.000	-.607	.607
	Glass's delta	.224	.000	-.620	.620

a. The denominator is used to estimate the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

The results of the effect size analysis conducted on Speech Act Theory indicate that although there are differences between ChatGPT and Gemini AI in applying It theory, these differences are not large enough to be considered significant. Based on Cohen's d, which measures the difference between the two groups using pooled standard deviation, the value obtained for Task 1 is 0.444, indicating a medium effect size. However, the Confidence Interval range that includes 0 (-0.620 to 0.620) suggests that the observed difference may not be consistent enough to be regarded as significant or practical. In Task 2, the Cohen's d value decreases to 0.224, indicating a small effect size. It further supports the conclusion that although there is a difference, it is not substantial enough to affect the results significantly.

Hedges' correction, which also uses pooled standard deviation and a correction for small sample sizes, yields results similar to Cohen's d. The Hedges' correction value for Task 1 is 0.453, indicating a medium effect size. However, as with Cohen's d, the Confidence Interval for Hedges' correction also includes 0 (-0.607 to 0.607), suggesting that the difference between the two models is still not statistically significant. In Task 2, the Hedges' correction value becomes 0.228, indicating a small effect size with a Confidence Interval that includes 0 (-0.607 to 0.607), further emphasizing that the observed difference between the two models is insignificant enough to be considered important.

Glass's delta, which uses the sample standard deviation of the control group (Gemini AI), provides results similar to Cohen's d and Hedges' correction. The Glass's delta value for Task 1 is 0.444, still indicating a medium effect size, but the Confidence Interval range that includes 0 (-0.620 to 0.620) suggests significant uncertainty in Its result. For Task 2, the Glass's delta value is 0.224, indicating a small effect size, with a Confidence Interval that includes 0 (-0.620 to 0.620), further affirming that the difference between the two models is not substantial enough to be considered significant or practically relevant.

Although both models show differences in the application of Speech Act Theory, the results indicate that these differences are not large enough to be considered significant or have a meaningful impact. The small to medium effect sizes and the confidence interval ranges that include 0 suggest that the differences between ChatGPT and Gemini AI in speech act theory are likely due to random variation in the data and are not strong enough to draw further conclusions.

Conclusion

Both AI models, ChatGPT and Gemini AI, are similar in applying pragmatic principles, particularly in Grice's Maxims. Both models successfully maintain relevance, quantity, quality, and manner in everyday conversation with high consistency. These results indicate that Grice's Maxims is a highly effective pragmatic theory for assessing the ability of AI models to generate coherent and relevant dialogue. Although there are slight differences between the two models, these differences are not substantial enough to alter the conclusion that both can efficiently apply these fundamental principles.

The application of Speech Act Theory presents more significant challenges, with both models demonstrating limitations in recognizing and responding to deeper communicative intentions, such as illocution and perlocution. The lower values for Speech Act Theory suggest that while ChatGPT and Gemini AI are adept at identifying locutionary acts, they struggle to comprehend the impact or intent behind these statements. It indicates that while both models can maintain clear and relevant communication, further development is needed to handle more complex and intention-based conversations more effectively.

Bibliography

Austin, J. L. (1975). *How to do things with words*. Harvard university press.

Beikian, A. (2024). A Critical Analysis of Mirza Reza Kermani's Interrogation Through Grice's Conversational Maxims with the Assistance of Artificial Intelligence (AI). *Critical Applied Linguistics Studies*, *1*(1), 161-182. 10.22034/jcals.1.1.161

Boxer, D. (1996). Ethnographic interviewing as a research tool in speech act analysis: The case of complaints. *Speech acts across cultures: Challenges to communication in a second language*, 217-239. https://doi.org/10.1515/9783110219289

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

Dornberger, R., Inglese, T., Korkut, S., & Zhong, V. J. (2018). Digitalization: Yesterday, today and tomorrow. *Business Information Systems and Technology 4.0: New Trends in the Age of Digital Change*, 1-11. https://doi.org/10.1007/978-3-319-74322-6_1

Ephratt, M. (2012). "We try harder"–Silence and Grice's cooperative principle, maxims and implicatures. *Language & Communication*, *32*(1), 62-79. https://doi.org/10.1016/j.langcom.2011.09.001

Firdaus, T., Sholeha, S. A., Jannah, M., & Setiawan, A. R. (2024). Comparison of ChatGPT and Gemini AI in Answering Higher-Order Thinking Skill Biology Questions: Accuracy and Evaluation. *International Journal of Science Education and Teaching*, *3*(3), 126-138. https://doi.org/10.14456/ijset.2024.11

Grice, H. P. (1990). 1975 Logic and Conversation. *The Philosophy of Language*.

Kaas, M. H., & Habli, I. (2024). Assuring AI safety: fallible knowledge and the Gricean maxims. *AI and Ethics*, 1-14. https://doi.org/10.1007/s43681-024-00490-x

Kasap, S., & Pashayeva, S. (2020). Pragmatic Competence And Foreign Language Teaching. S. Kasap İçinde, *Applied Linguistics İn Elt Classrooms* (S. 69). Ankara: Anı Yayıncılık.

Kasap, S & Dağdemir, B. (2021). Pragmatics-A Glance at the Cooperative Principle and Grice's Maxims. *Eurasian Confrence on Language & Social Science*.

Kasirzadeh, A., & Gabriel, I. (2023). In conversation with artificial intelligence: aligning language models with human values. *Philosophy & Technology*, *36*(2), 27. https://doi.org/10.1007/s13347-023-00606-x

Lee, B. J. (2024). Exploring the potential of AI for pragmatics instruction. *Technology in Language Teaching & Learning*, *6*(3), 1521-1521.https://doi.org/10.29140/tltl.v6n3.1521

Lee, Q., Lavingia, K. S., & Amendola, M. (2024). The performance of large language models on fictional consult queries indicates favorable potential for AI-assisted vascular surgery consult handling. *JVS-Vascular Insights*, *2*, 100052. https://doi.org/10.1016/j.jvsvi.2023.100052

Miller, J. H. (2001). *Speech acts in literature*. Stanford University Press.

Nam, Y., Chung, H., & Hong, U. (2023). Language Artificial Intelligences' Communicative Performance Quantified Through the Gricean Conversation Theory. *Cyberpsychology, Behavior, and Social Networking*, *26*(12), 919-923. https://doi.org/10.1089/cyber.2022.0356

Rane, N., Choudhary, S., & Rane, J. (2024). Gemini versus ChatGPT: applications, performance, architecture, capabilities, and implementation. *Performance, Architecture, Capabilities, and Implementation (February 13, 2024)*. https://dx.doi.org/10.2139/ssrn.4723687

Recanati, F. (1987). Literalness and other pragmatic principles. *Behavioral and Brain Sciences*, *10*(4), 729-730. https://doi.org/10.1017/S0140525X00055540

Saka, A. B., Oyedele, L. O., Akanbi, L. A., Ganiyu, S. A., Chan, D. W., & Bello, S. A. (2023). Conversational artificial intelligence in the AEC industry: A review of present status, challenges and opportunities. *Advanced Engineering Informatics*, 55, 101869. https://doi.org/10.1016/j.aei.2022.101869

Searle, J. R., Kiefer, F., & Bierwisch, M. (Eds.). (1980). *Speech act theory and pragmatics* (Vol. 10). Dordrecht: D. Reidel. https://doi.org/10.1007/978-94-009-8964-1

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

Sun, Y., Li, S., & Yu, L. (2022). The dark sides of AI personal assistant: effects of service failure on user continuance intention. *Electronic Markets*, *32*(1), 17-39. https://doi.org/10.1007/s12525-021-00483-2

World Bank. (2024). *Who on Earth Is Using Generative AI?*

Wu, S. (2022). Design of intelligent customer service questioning and answering a system for power business scenario based on ai technology. *Mathematical Problems in Engineering*, *2022*(1), 5279919. https://doi.org/10.1155/2022/5279919