

THE ROLE OF PRAGMATICS IN TRAINING RESPONSIBLE AI COMMUNICATION MODELS

Amende A. Charles,
Dalmeida, Chibuzor Ozoemena
&
Ikwulono, Rebecca Eyum

Abstract

This paper examines the role of pragmatics in training responsible AI communication models. As artificial intelligence systems take on increasingly conversational roles, the ability to interpret and generate meaning beyond the literal becomes essential. This paper examines how pragmatics, the study of language in context, contributes to the development of responsible AI communication models. Many AI systems succeed at processing syntax and semantics but struggle with contextual cues such as indirect requests, politeness strategies, cultural norms, and implied meanings. These gaps can lead to misunderstandings, insensitive responses, or communication that falls short of user expectations. This paper argues that embedding pragmatic principles into AI training pipelines strengthens an AI system's capacity to interpret user intent responsibly. Drawing from speech act theory, Gricean Maxims, politeness theory, and relevance theory, the discussion highlights how pragmatic awareness helps AI models adjust tone, recognize conversational boundaries, and avoid context-inappropriate output. Special attention is given to interactions involving young users and culturally diverse groups, where pragmatic missteps can have emotional or social consequences. Challenges such as modelling dynamic human context, preventing reinforcement of cultural bias, and ensuring ethical guardrails are also addressed. The paper concludes that integrating pragmatic reasoning is not simply an enhancement but a necessary condition for developing AI systems that communicate safely, sensitively, and ethically across digital environments.

Keywords: Pragmatics, Artificial Intelligence, Responsible AI, Language Models, AI Ethics, Human–Computer Interaction

Introduction

The proliferation of artificial intelligence (AI) systems in conversational and human-centred applications has intensified scholarly and practical interest in the communicative capacities of these systems. Historically, language processing systems focused on syntax and semantics, the formal structure and literal meaning of language, respectively, but increasingly, researchers recognize that this is insufficient for nuanced and contextually appropriate communication (Sato et al., 2025). Pragmatics, the branch of linguistics that studies how context and social factors shape meaning beyond literal interpretation, addresses precisely the communicative dimensions where many AI systems still struggle (Levinson, 2000; Sato et al., 2025). The seminar topic of this paper, *The Role of Pragmatics in Training Responsible AI Communication Models*, situates this need within a broader call for AI that is both effective and responsible in its interaction with diverse users.

Pragmatic competence enables humans to interpret indirect requests, recognize implied meanings, negotiate politeness expectations, and adjust utterances according to cultural norms and social goals. These capabilities are essential in everyday discourse and are increasingly expected from AI systems that are designed to interact with humans across a variety of contexts, such as customer service, education, health information, and personal assistants. Yet contemporary natural language processing (NLP) and large language models (LLMs), even when achieving high accuracy in syntax and semantics, frequently falter at pragmatic reasoning. They may misinterpret indirect speech, fail to recognize conversational implicatures, or generate responses that are context-inappropriate or insensitive (Eragamreddy, 2025). These gaps can lead to

misunderstandings, reduced user satisfaction, and, in sensitive contexts, emotional or social harm.

Bridging these gaps requires embedding pragmatic principles directly into AI training and evaluation pipelines. While conventional training focuses on pattern recognition from large textual corpora, emerging research demonstrates that supplying models with explicit representations or prompts grounded in pragmatic theory can improve their interpretation of implied meanings and contextually grounded utterances (Sato et al., 2025). For example, methods that integrate Gricean maxims and relevance principles into prompts have enhanced performance on tasks requiring pragmatic inference, suggesting that pragmatic frameworks provide valuable scaffolding for reasoning beyond literal text (Sato et al., 2025).

Integrating pragmatics into AI development aligns with ethical and responsible AI imperatives. Responsible communication models should not only produce syntactically valid responses but also demonstrate contextual sensitivity, respect for user norms, and awareness of conversational boundaries, particularly with young users or culturally diverse populations, where miscommunication can have amplified consequences. Scholars argue that without attention to pragmatics, AI systems risk reinforcing cultural biases or delivering unsafe outputs that conflict with user expectations (Eragamreddy, 2025). Thus, pragmatic training is not an optional enhancement but a necessary foundation for communication models that are safe, sensitive, and socially responsible.

This paper explores several foundational pragmatic theories and their relevance to AI communication: Gricean maxims, Speech Act Theory, Politeness Theory, and Relevance Theory. Gricean maxims provide a cooperative framework for effective conversation, addressing contextual norms of quantity, quality, relation, and manner. Speech Act Theory conceptualizes utterances as actions asserting, promising, and requesting, grounding the performative aspects of language. Politeness Theory illustrates how social goals like face-saving shape communicative choices. Relevance Theory emphasizes how speakers

and listeners negotiate meaning through cognitive principles of relevance and inference. Together, these theories illuminate facets of language use that are especially salient for AI systems that operate beyond the literal.

The remainder of the paper reviews these theories, discusses challenges in modelling human-like pragmatics in AI, and examines approaches to embedding pragmatic reasoning into training pipelines. By drawing on recent research in computational linguistics and machine learning, this study demonstrates how pragmatics strengthens AI communication models, making them more context-aware, culturally attuned, and ethically grounded.

Conceptual Review

Pragmatics

Pragmatics is the branch of linguistics concerned with the study of language in context. Unlike syntax and semantics, which focus on the structure and literal meaning of words, pragmatics deals with how meaning is constructed and interpreted based on social context, speaker intention, cultural background, and shared knowledge. Yule (2020) defines pragmatics as “the study of speaker meaning,” emphasizing that interpretation often transcends what is explicitly said. This includes interpreting indirect requests, sarcasm, presuppositions, and implicatures.

In contemporary discourse, pragmatics is foundational for understanding the richness of human interaction. According to Huang (2021), pragmatics enables interlocutors to decode meaning that is implied but not overtly expressed. For example, the question “Can you pass the salt?” is often a request, not a literal inquiry about capability. Recognizing such intentions is critical in both human and AI conversations.

When applied to AI development, pragmatics becomes indispensable. While many AI models are adept at syntax and

semantics, they often fail to grasp the nuanced intent behind user utterances (Sato et al., 2025). Without pragmatic competence, AI systems may misinterpret indirect language, fail to adhere to politeness norms, or generate socially inappropriate responses. Levinson (2022) emphasizes that pragmatics provides the cognitive scaffolding required for responsible and sensitive interaction, especially in multicultural or high-stakes communication environments. Hence, pragmatics is not a peripheral concern but a core requirement for any AI model designed for conversational tasks. Embedding pragmatic reasoning allows machines to engage more naturally with humans, aligning output with user expectations and cultural appropriateness (Eragamreddy, 2025).

Artificial Intelligence (AI)

Artificial Intelligence (AI) refers to the field of computer science concerned with building machines and systems capable of performing tasks that normally require human intelligence. These tasks include problem-solving, learning, perception, natural language understanding, and decision-making. Russell and Norvig (2020) describe AI as “the study of agents that perceive their environment and take actions that maximize their chances of achieving goals.” Contemporary definitions of AI often emphasize autonomy, adaptability, and data-driven reasoning.

Since the late 2010s, the development of AI has accelerated, particularly in the domain of machine learning and natural language processing (NLP). Large language models (LLMs), such as GPT and BERT, have transformed how machines interpret, generate, and respond to human language (Vaswani et al., 2017; Devlin et al., 2018). These systems are trained on massive textual datasets and rely on statistical relationships between words to produce coherent output. However, this statistical grounding alone does not guarantee communicative appropriateness.

The rise of conversational AI (e.g., chatbots, virtual assistants, customer service bots) has increased the urgency for models that not only understand words but also grasp their pragmatic implications.

According to Abdulrazzak (2025), current models often falter in context-rich or emotionally sensitive interactions because they lack real-world grounding and pragmatic insight. They may generate accurate but inappropriate responses, leading to user dissatisfaction or even harm. Therefore, responsible AI development requires the integration of ethical, contextual, and pragmatic awareness. AI systems must be designed to interpret meaning not just in terms of data but also in terms of human goals, emotions, and cultural expectations (Sato et al., 2025). AI without pragmatics is limited to mechanical interaction; AI with pragmatics moves closer to truly intelligent and ethical communication.

Communication Models

Communication models are theoretical frameworks that describe the process through which information is transmitted from a sender to a receiver. Traditional models, such as Shannon and Weaver's (1949) transmission model, focused on linear information exchange encoding, sending, and decoding but did not account for context, feedback, or non-verbal cues. Later developments, such as Schramm's interactional model and Barnlund's transactional model, introduced feedback loops and emphasized the dynamic nature of communication.

In contemporary communication theory, especially digital communication, models increasingly focus on interactivity, shared context, and adaptation. West and Turner (2019) note that modern communication models must consider social, cultural, and psychological factors that influence how messages are constructed and interpreted. These models go beyond the literal content of communication to include the roles of tone, context, and relationship dynamics.

For AI systems designed to engage in human-like communication, traditional transmission models are insufficient. AI must be able to simulate both the encoding and interpretation of messages based on contextual knowledge and user intent. This is where

communication models informed by pragmatics become essential (Leech & Kasper, 2021). In AI-human interactions, meaning is not static; it is co-constructed, negotiated, and influenced by context.

According to Eragamreddy (2025), communication models used in AI must support dynamic turn-taking, topic tracking, and sensitivity to indirectness, especially in multilingual or multicultural contexts. Without an embedded communication model, AI systems risk producing linear, decontextualized responses that do not meet users' social or emotional expectations. Thus, communication models serve as a blueprint for developing AI systems capable of meaningful, adaptive, and responsible interaction. Incorporating pragmatics into these models enables machines to transcend mere data exchange and participate in authentic communicative acts.

Methodology

This study adopts a qualitative, exploratory research design to examine the role of pragmatics in training responsible AI communication models. Given the theoretical and interdisciplinary nature of the study, drawing from linguistics, artificial intelligence, and ethics, a qualitative approach is most appropriate for analysing concepts, theoretical frameworks, and existing scholarly contributions.

The research employs a conceptual and analytical approach, focusing on the interpretation and synthesis of existing literature rather than on empirical data collection. This design enables an in-depth exploration of how pragmatic theories can be integrated into AI training processes and how they contribute to the development of responsible communication models.

Data for this study are derived from secondary sources, including peer-reviewed journal articles in pragmatics, computational linguistics, and AI ethics; scholarly books and academic publications; conference proceedings on natural language processing (NLP) and AI communication; and recent studies on large language models (LLMs) and conversational AI systems. These sources were selected based on

their relevance, credibility, and their contribution to the intersection of pragmatics and AI development.

The study employs thematic content analysis to systematically examine the selected literature. This method allows for the identification, organisation, and interpretation of recurring patterns and ideas relevant to the research focus. The analytical process is carried out in four stages. First, key themes such as context awareness, implicature, politeness, and intention recognition are identified from the literature. Second, these themes are categorised into broader pragmatic components and AI communication challenges. Third, the study interprets how these themes relate to responsible AI communication and model training. Finally, insights from various sources are synthesised to propose strategies for embedding pragmatic principles into AI systems.

The choice of a qualitative approach is justified on several grounds. Pragmatics, as a field, is primarily concerned with context, meaning, and interpretation, which are best examined through conceptual and interpretive analysis rather than quantitative measurement. Furthermore, the study seeks to bridge theoretical insights and practical applications in AI development, rather than to test specific algorithms or models. Finally, the availability of extensive and relevant literature provides sufficient depth for critical analysis and synthesis, making a qualitative methodology both appropriate and effective for achieving the objectives of this study.

Data Presentation and Analysis

This study presents and analyses data qualitatively, drawing on themes identified from the reviewed literature on pragmatics and AI communication. The presentation is organised around key pragmatic components and AI communication challenges, followed by an interpretation of findings using established pragmatic theories.

Presentation of Pragmatic Components in AI Communication

The analysis of the literature reveals five major pragmatic components that influence the performance of AI communication models.

Context Awareness

The data indicate that context awareness remains a significant limitation in current AI systems. Although AI models can process large amounts of textual data, they often fail to incorporate situational, social, and cultural context into their responses. As a result, outputs may be grammatically correct but pragmatically inappropriate. This finding suggests that context modelling is essential for responsible AI communication.

Speaker Intention Recognition

The findings show that AI systems frequently struggle to identify speaker intention, particularly in indirect speech. For instance, utterances that function as requests, warnings, or emotional expressions are often interpreted literally. This limitation reduces the effectiveness of AI in real-world communication, where meaning is often implied rather than explicitly stated.

Implicature and Implied Meaning

Another key finding is the inability of AI systems to interpret implicature. The data reveal that AI models tend to prioritise literal meanings while ignoring implied intentions, sarcasm, or indirect cues. This leads to misunderstandings and responses that fail to align with user expectations.

Politeness and Face Management

The analysis shows that AI systems often lack sensitivity to politeness strategies and face-saving mechanisms. Responses generated by AI may appear blunt, overly direct, or culturally inappropriate. This highlights the importance of integrating politeness theory into AI training to ensure socially acceptable interactions.

Deixis and Context-Dependent Expressions

The findings also reveal challenges in handling deixis, including references to time, place, and participants. AI systems sometimes misinterpret such expressions due to insufficient context tracking, resulting in ambiguous or confusing responses.

Presentation of AI Communication Challenges

The data further reveal several overarching challenges affecting AI communication models.

Context Misinterpretation

AI systems often misinterpret the broader context of interactions, particularly in complex or sensitive scenarios. This limitation underscores the gap between linguistic accuracy and communicative appropriateness.

Lack of Intention Recognition

The inability to infer communicative intent remains a persistent issue. AI models rely heavily on surface-level patterns, which limit their ability to understand nuanced human communication.

Insensitive and Culturally Inappropriate Responses

The analysis indicates that AI systems frequently fail to adapt to cultural and social norms. This results in responses that may be perceived as insensitive or inappropriate, especially in multicultural environments.

Analysis Based on Pragmatic Theories

The findings are further analysed using key pragmatic frameworks:

1. Gricean Maxims: The data show that AI systems often violate the maxims of relevance, quantity, and manner, producing responses that are either insufficient, excessive, or unclear.

2. Speech Act Theory: AI models demonstrate limited ability to recognize and perform appropriate speech acts, particularly in indirect or emotionally nuanced communication.
3. Politeness Theory: The findings reveal inadequate management of face needs, resulting in responses that may threaten user dignity or social expectations.
4. Relevance Theory: AI systems struggle to identify the most contextually relevant interpretations, especially when meaning is implicit.

Integrated Analysis of Findings

The overall analysis reveals that the primary limitations of AI communication models are pragmatic rather than structural. While these systems demonstrate high levels of syntactic and semantic competence, they cannot interpret language as a socially embedded phenomenon.

The findings further indicate that:

- (a) Effective communication requires more than linguistic accuracy; it requires contextual and social understanding
- (b) AI systems must be trained to recognize user intention, implied meaning, and cultural norms
- (c) Pragmatic failures can lead to miscommunication, user dissatisfaction, and ethical risks

Summary of Data Presentation and Analysis

The table below summarizes the key findings:

Pragmatic Element	Observed AI Limitation	Analytical Insight
Context Awareness	Misinterpretation of situations	Need for context-aware models
Intention Recognition	Literal interpretation of indirect speech	Integration of speech act modelling is required
Implicature	Failure to detect implied meaning	Enhancement of inferential reasoning is needed

Pragmatic Element	Observed AI Limitation	Analytical Insight
Politeness	Inappropriate tone and style	Inclusion of cultural and social norms necessary
Deixis	Misunderstanding of references	Improved context tracking is required
Ethics	Biased or harmful outputs	Stronger ethical and pragmatic alignment is needed

Components of Pragmatics

The key components of pragmatics can be grouped into several core phenomena directly relevant to training responsible AI communication models: contextual meaning, speaker intention, implicature, politeness and face management, and deixis. Key pragmatics components include: Contextual meaning, Speaker intention, Implicature, Politeness and face management and Deixis.

Contextual meaning

Contextual meaning concerns how utterances are interpreted in light of situational, social, and cultural factors, including who is speaking, to whom, and for what purpose. Meaning is co-constructed through shared knowledge and discourse history rather than derived solely from word definitions (Tasheva, 2025; Noveck & Reboul, 2008). In AI systems, failure to model context can result in responses that are technically accurate yet emotionally or socially inappropriate, particularly in sensitive situations. Studies show that large language models still struggle when interpretation depends on subtle contextual cues rather than surface text (Settaluri et al., 2024).

Speaker's intention

Contextual meaning concerns how utterances are interpreted considering situational, social, and cultural factors, including who is

speaking, to whom, and for what purpose. Meaning is co-constructed through shared knowledge and discourse history rather than derived solely from word definitions (Tasheva, 2025; Noveck & Reboul, 2008). In AI systems, failure to model context can result in responses that are technically accurate yet emotionally or socially inappropriate, particularly in sensitive situations. Studies show that large language models still struggle when interpretation depends on subtle contextual cues rather than surface text (Settaluri et al., 2024).

Implicature

Implicature involves meanings that are implied rather than explicitly stated. Humans rely on implicature to interpret hints, understatements, and indirect requests (Eragamreddy, 2024; Siregar, 2024). AI models frequently fail in this area, responding to literal content while missing underlying intent, a gap highlighted by pragmatics-oriented evaluation benchmarks (Settaluri et al., 2024). Such failures can be harmful in high-stakes contexts.

Politeness and face management

Politeness and face management address how speakers maintain social dignity and relationships through linguistic choices. These strategies are culturally dependent and crucial for respectful interaction (Siregar, 2024; Tasheva, 2025). AI systems that underuse politeness markers or mishandle face-threatening acts risk sounding rude or insensitive, undermining trust (Karabulut, 2025).

Deixis

Deixis involves context-dependent expressions whose meaning depends on time, place, and participant roles. Misinterpreting deixis can cause confusion in dialogue, and benchmarks show AI still relies heavily on superficial cues rather than robust context tracking (Settaluri et al., 2024).

Overall, human communication depends on the interaction of these pragmatic components. Without them, AI may produce fluent

responses that are socially or ethically problematic. Embedding these components into AI training and evaluation is therefore essential for developing responsible, context-sensitive, and culturally aligned communication models.

Pragmatics and Responsible AI Communication

Responsible AI communication involves designing AI systems whose language use is ethical, fair, safe, transparent, and aligned with human values. Beyond technical accuracy, AI must avoid pragmatic failures that harm users or erode trust. Pragmatics provides the tools for achieving this by focusing on context, intention, implicature, and social relations.

Context Awareness

Context awareness enables AI systems to interpret and generate responses based on situational, social, and cognitive factors such as user role, interactional setting, and potential consequences. (Schnell, 2021). Without this capacity, AI may provide factually correct but harmful or inappropriate responses, especially in high-risk domains like medicine or law. Embedding pragmatic context sensitivity allows AI to adjust tone, detail, and caution, directly supporting safety, responsibility, and non-maleficence in communication (Bunt, 2023)

Interpretation of Implicit Meaning

Human communication relies heavily on implicit meaning, including implicature, indirect speech acts, irony, and sarcasm. AI systems often struggle to interpret such meanings, leading to missed distress signals or inappropriate replies. Recent studies show that large language models still struggle with complex cases of implicit disagreement, irony, and sarcasm, despite strong performance on surface linguistic tasks (Shulginov, 2025; Sarcasm Bench, 2026). Failure

to handle implicit meaning can result in serious risks, such as ignoring veiled requests for help.

Politeness and Social Norms

Politeness and face management are classic themes in pragmatics and play a decisive role in how communication is experienced as respectful, supportive, or humiliating. Bowman (2024) shows that the degree and type of politeness employed by chatbots significantly shape user experience, affecting perceptions of care, trustworthiness, and emotional support. However, politeness can also be overdone or mis-calibrated, coming across as condescending or insincere, which further illustrates the need for nuanced pragmatic modelling rather than simple rule-based politeness filters.

Ethical Alignment and Harm Reduction

Finally, pragmatics supports ethical alignment and harm reduction by enabling AI systems to recognise the intentional and interactional profile of user inputs. AI ethics in communication highlights risks such as manipulation, disinformation, emotional exploitation, and the reinforcement of harmful stereotypes (Bowen, 2024; “AI Ethics in Communication,” 2025). To mitigate these risks, AI must distinguish between benign and harmful requests, detect attempts to solicit unethical assistance, and respond in ways that uphold safety and public values.

Discussion of Findings

The findings reveal that the major weaknesses of contemporary AI communication models are fundamentally pragmatic rather than syntactic or semantic. Although these systems demonstrate high levels of linguistic fluency, they frequently misinterpret context, fail to recognize user intentions, generate insensitive or culturally inappropriate responses, and sometimes produce ethically problematic content. These shortcomings stem from the inability of current models to treat language as social action embedded in context. When examined

through key pragmatic theories, Speech Act Theory, Gricean Maxims, Politeness Theory, and Relevance Theory, the findings consistently indicate that responsible AI communication depends on the explicit integration of pragmatic reasoning into model training and deployment.

Speech Act Theory explains why AI systems often perform poorly in emotionally sensitive or high-stakes interactions. Human utterances function as actions, such as requesting, warning, or comforting, yet AI models typically focus on literal meaning while failing to capture illocutionary force. This limitation becomes evident in the difficulty models have in recognizing indirect requests, expressions of distress, or subtle refusals. Evidence that pragmatically annotated data and human-in-the-loop evaluation improve response appropriateness confirms that modelling speech acts enhances contextual sensitivity and responsible behaviour.

The findings also align with Grice's Cooperative Principle. Issues such as irrelevant, overly verbose, or misleading responses reflect systematic violations of the maxims of Relation, Quantity, and Quality. Improvements observed with context-aware architectures suggest that incorporating discourse and user goals helps approximate cooperative conversational behaviour.

Politeness Theory sheds light on negative user reactions to blunt or culturally insensitive AI responses. AI systems often fail to manage face needs appropriately, leading to ethical lapses that function as unmitigated face-threatening acts. Training with culturally diverse data and human feedback improves trust and respect, reinforcing the theory's emphasis on culture-specific politeness norms.

Finally, Relevance Theory explains AI difficulties with implicature, sarcasm, and indirect meaning. Emerging techniques such as context-aware models and pragmatic benchmarks represent attempts to approximate relevance-guided inference. Overall, the findings confirm that pragmatic competence is essential for developing

responsible, context-sensitive, and ethically aligned AI communication models.

Conclusion

This seminar has demonstrated that the primary weaknesses of contemporary AI communication models are fundamentally pragmatic rather than syntactic or semantic. Although large language models exhibit impressive fluency, they frequently misinterpret context, fail to recognise user intentions, mishandle politeness and cultural norms, and sometimes generate ethically problematic responses. These shortcomings are best understood through pragmatic theories such as Speech Act Theory, Gricean Maxims, Politeness Theory, and Relevance Theory, all of which emphasize language as situated social action rather than decontextualized text. The study's conceptual clarifications showed that responsible AI communication requires integrating not only linguistic structure and meaning but also social functions, intentions, and contextual constraints. Analysis of key pragmatic components, contextual meaning, speaker intention, implicature, politeness, and deixis, revealed that human communication depends on these elements to avoid misunderstanding and manage relationships, particularly in sensitive domains such as healthcare, education, and counselling. The examination of AI communication models and their limitations further confirmed that statistical language learning alone cannot guarantee responsible interaction. Consequently, explicit pragmatic modelling emerges as a necessary condition for responsible AI, supported through pragmatically annotated data, context-aware architectures, human-in-the-loop evaluation, and cultural adaptation. While challenges such as contextual complexity, subjective interpretation, data scarcity, and computational demands persist, these do not diminish the relevance of pragmatics.

References

- Abbas, M. (2025). Pragmatics and responsible conversational agents in higher education. *Pakistan Social Sciences Review*, 9(1), 1–20.
- Aijmer, K. (2018). *How to do corpus pragmatics on pragmatically annotated data: Speech acts and beyond*. John Benjamins.
- Akbar, A. (2024). Large language models: Challenges and opportunities. *Ethic AI Journal of Responsible Technology*, 2(1), 15–32.
- Artsi, A., & colleagues. (2025). Pragmatic risks of LLMs in clinical communication. *Journal of Medical Systems*, 49(2), 1–14.
- Aziz, A. (2025). Pragmatic failures in AI-mediated communication. *International Journal of Language and Communication Studies*, 7(1), 45–63.
- Bowen, S. A. (2024). AI ethics in communication: From abstract principles to design-level practice. *Public Relations Review*, 50(2), 1–11.
- Bunt, H., & Petukhova, V. (2023). Semantic and pragmatic precision in conversational AI systems. *Frontiers in Artificial Intelligence*, 6, 896729.
- Chang, S., Kim, J., & Lee, H. (2020). Context-aware transformer transducer for dialogue modelling. *Proceedings of Interspeech 2020*, 1–5.
- Chen, L. (2024). Pragmatics-aware natural language processing: A research agenda. *Journal of Pragmatics*, 214, 1–16.
- De Felice, R. (2024). Pragmatically annotated corpora and speech function analysis. *Italian Journal of Computational Linguistics*, 10(1), 493–507.
- de Souza Monteiro, T. (2024). Conversational-linguistic politeness parameters for chatbot design. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1–14).

- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT 2019* (pp. 4171–4186).
- Dey, R. (2023). Primary characteristics of English pragmatics in applied linguistics. *International Journal of English Linguistics*, 13(2), 15–28.
- Eragamreddy, S. (2024). Exploring pragmatics: Uncovering the layers of language and meaning. *International Journal of Current Science Research and Review*, 7(3), 3277–3284.
- Eragamreddy, S. (2025). Pragmatic competence in human-AI interaction. *ESP Today*, 13(1), 55–73.
- Fried, D., et al. (2022). Pragmatics in neural language models. In *Proceedings of EMNLP 2022* (pp. 800–815).
- Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: Is ChatGPT a blessing or a curse? *Frontiers in Education*, 8, 1166682.
- Huang, Y. (2017). *The Oxford handbook of pragmatics*. Oxford University Press.
- Karabulut, A. (2025). Style variation and politeness in LLM-based chatbots. *American International Journal of Multidisciplinary Scientific Research*, 7(1), 1–15.
- Laakso, V. (2024). Social and ethical issues of large language models. *Proceedings of the Nordic Conference on Human-Computer Interaction*, 1–12.
- Levinson, S. C. (2022). Pragmatics and the challenge of conversational AI. *Annual Review of Linguistics*, 8, 1–20.
- Liu, X., Zhang, Y., & Li, J. (2025). Culture alignment of large language models through simulated social interactions. *Proceedings of ACL 2025*, 2230–2245.
- Luo, J. (2025). A cross-cultural pragmatic analysis of human-AI politeness strategies. *Pakistan Academy of Social Sciences Review*, 3(1), 1–14.

- Ma, B., Li, Y., Zhou, W., Gong, Z., Liu, Y. J., Jasinskaja, K., Friedrich, A., Hirschberg, J., Kreuter, F., & Plank, B. (2025). Pragmatics in the era of large language models: A survey on datasets, evaluation, opportunities, and challenges. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics* (pp. 7600–7624).
- Memarian, B. (2024). Human-in-the-loop machine learning for safe AI systems. *AI Safety Journal*, 1(2), 30–52.
- Monrad, M. (2024). Chatbots, emotions, and the reshaping of feeling rules. *Emotions and Society*, 6(2), 243–260.
- Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2021). Ethics as a service: A pragmatic operationalisation of AI ethics. *Minds and Machines*, 31(2), 239–256.
- Naji, A. (2025). Integrating pragmatic principles into virtual assistants. *Canadian Journal of Linguistics and Pragmatic Communication*, 7(1), 1–18.
- Naidoo, S. (2025). A lifecycle framework for culturally responsive AI chatbots (CRAIF-C). *International Journal of Human-Computer Studies*, 186, 103148.
- Noveck, I. A., & Reboul, A. (2008). Experimental pragmatics: A Gricean turn in the study of language. *Trends in Cognitive Sciences*, 12(11), 425–431.
- Orsini, M. (2023). Indirect speech acts in large language models: A pragmatic evaluation [Master’s thesis, University of Padua].
- Pooja, R. (2025). Context-aware dialogue generation with memory-augmented transformers. *Proceedings of ICAART 2025*, 312–321.
- Qin, J. (2025). Opportunities and challenges for large language models in primary healthcare. *npj Digital Medicine*, 8(1), 1–12.

- Ruane, F., & colleagues. (2019). Social and ethical issues in conversational AI. In *Proceedings of the 30th Irish Conference on Artificial Intelligence and Cognitive Science* (pp. 1–12).
- Russell, S., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
- Sato, M., Tanaka, Y., & Kobayashi, H. (2025). Prompting large language models with pragmatic theory. *Journal of Natural Language Engineering*, 31(1), 55–79.
- Schnell, M. (2021). Context, cognition, and meaning: A pragmatic view. *Journal of Pragmatics*, 180, 50–63.
- Settaluri, K., et al. (2024). PragmatiBench: Evaluating pragmatic reasoning in large language models. In *Proceedings of ACL 2024* (pp. 5100–5115).
- Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press.
- Shulginov, V., et al. (2025). Irony and sarcasm understanding in large language models. In *Proceedings of Dialogue 2025* (pp. 350–365).
- Siregar, N. (2024). Politeness strategies and pragmatic competence in digital discourse. *Journal of Innovative Language Learning*, 6(1), 90–106.
- Smaglii, L. (2025). Cross-cultural pragmatics in English as a lingua franca. *Arab World English Journal*, 16(4), 310–330.
- Ta, M. (2025). Content moderation and large language models: A critical study. *Information, Communication & Society*, 28(6), 1–20.
- Tasheva, D. (2025). The role of context in pragmatic interpretation. *Science and Research Journal*, 9(2), 215–222.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems 30* (pp. 5998–6008).
- West, R., & Turner, L. (2019). *Introducing communication theory: Analysis and application* (7th ed.). McGraw-Hill.

- Yakasai, M. S. (2025). Exploring pragmatic implications of artificial intelligence on learning and authorship. *Journal of the Linguistic Association of Nigeria*, 28(1), 173–192.
- Yu, D., Li, L., Su, H., & Fuoli, M. (2023). Assessing the potential of AI-assisted pragmatic annotation: The case of apologies. *arXiv Preprint, arXiv:2305.08339*.
- Yule, G. (2020). *The study of language* (7th ed.). Cambridge University Press.