

---

# CoRe Essay 5

## Communication

---

**Yuntian Gu**  
Yuanpei College  
Peking University  
guyuntian@stu.pku.edu.cn

### Abstract

The distinction between language imitation and genuine communication in AI has significant implications for the development of sophisticated language and chat robots. This essay proposes a new evaluative framework for assessing language models' capacity for true communication, characterized by understanding intent, emotional depth, and contextual adaptability. We introduce specific tasks to measure models' ability to decipher intent, maintain contextual coherence, and adapt to varied dialogue scenarios. Additionally, we discuss the integration of nonverbal communication in chat robots, focusing on facial expressions, gestures, and prosodic cues to enrich and clarify verbal interactions. By demonstrating the use of context-sensitive nonverbal behavior, we aim to establish a more human-like, efficient, and emotionally congruent communicative process in AI. This research represents a step toward enabling AI systems to engage in more meaningful and collaborative exchanges with humans, advancing them beyond mere textual response generators to entities capable of understanding and participating in the full spectrum of human communication.

## 1 Introduction

The advent of sophisticated language models has revolutionized the way machines understand and generate human-like text. However, the distinction between simple imitation and genuine communication remains a contentious issue. The unique nuances of human interaction encapsulate more than the mere exchange of words; they involve a complex tapestry of intent, emotion, and adaptability [4, 1]. This paper seeks to establish rigorous evaluation frameworks that go beyond surface-level performance to decipher whether language models are truly communicating or merely mimicking. We propose tasks that challenge models to interpret the underlying intent of queries and exhibit adaptive communication across diverse contexts, which is a hallmark of genuine understanding.

Furthermore, in the realm of embodied artificial intelligence, the interplay between verbal and nonverbal communication is pivotal. To create an embodied chat robot that transcends the limitations of text-based models, it is essential to integrate nonverbal expressions that complement and enhance verbal interactions. This paper will explore the design of context-sensitive nonverbal cues, such as gestures, facial expressions, and prosodic elements, to create a more cohesive and intuitive communicative experience. We will examine how and when a chat robot should employ these nonverbal expressions in conjunction with verbal communication to effectively mirror the sophistication of human interaction.

## 2 Distinguishing simple imitation and true communication of Language Model

The sophistication of language models has traditionally been evaluated by their ability to generate text that is indistinguishable from that produced by humans. However, this imitation game often

overlooks the essence of true communication: the exchange of meaning, intention, and understanding that adapts dynamically to the changing contexts of dialogue. To distinguish between simple imitation and true communication, an evaluation framework must be designed with tasks that test beyond the language model's ability to produce plausible-sounding language. Below I provide three kinds of different tasks.

**Intentionality task:** this task delves into the model's proficiency in recognizing and responding to the intent underlying user queries, transcending the act of producing contextually relevant responses. For example, a task may involve a language model deciphering ambiguous requests or following complex, multi-step user instructions, mirroring real-life scenarios where communicative intent is often implicit.

**Adaptability Task:** true communication demands flexibility, necessitating language models to seamlessly transfer learning and apply knowledge across various domains and contexts. Adaptability tasks test this aspect by requiring models to apply their understanding in unfamiliar territories, such as switching between different genres of writing, social norms, or technical jargon, demanding an agile recalibration of language use.

**Context-awareness task:** this task is pivotal as they measure a language model's ability to maintain and apply context over a series of interactions. It assesses the language model's depth in processing and retaining information throughout a dialogue, rather than evaluating single, isolated exchanges. Such tasks could include maintaining thread coherence over extended conversations, with changing subjects and participants, requiring a sophisticated understanding of previous exchanges.

In summary, distinguishing between imitation and communication in language models necessitates a framework where the acquisition of language is an active, interactive process reflecting genuine understanding. It is within such frameworks that we can begin to engineer AI systems capable of engaging with humans in truly collaborative and communicative endeavors.

### **3 Chat Robots with nonverbal language**

The integration of nonverbal language into chat robots represents a paradigm shift from text-based interactions to more holistic, embodied communication. Nonverbal cues play an instrumental role in human communication, often conveying messages more succinctly and poignantly than words alone. The development of a chat robot with nonverbal capabilities involves the synthesis of these silent communicative forms to enrich the interaction between humans and machines [3, 2].

To mimic the depth of human interactions, it is critical for a chat robot to express a wide range of emotions through facial expressions. These expressions bolster the verbal communication by providing clear emotional context, enhancing the robot's ability to engage in empathetic and responsive dialogue. For instance, the presence of a frown or smile can substantially impact the perceived sincerity and warmth of the robot's responses.

Besides facial expression, gestures offer a rich lexicon of nonverbal communication that a chat robot can exploit to augment and clarify verbal messages. The design of gestures in chat robots should prioritize natural, human-like movements that can be easily recognized and interpreted by the user. By gesturing to indicate direction, display urgency, or complement a verbal instruction, robots can provide a more immersive and effective communication experience.

Equipped with many kinds of communication tools, the first and the most important behavior for chat robots is contextually relevant, ensuring that the use of gestures, facial expressions, and other nonverbal cues align with the accompanying verbal messages. The ability to discern when to use a nod, a shake of the head, or a shrug in concert with verbal communication is essential for coherence and mutual understanding.

In situations where brevity is valued, nonverbal language can substantially increase information efficiency. The strategic use of nonverbal cues by a chat robot can make complex interactions more intuitive and less time-consuming for users. Pointing to an object while naming it, for instance, is a prime example of combining verbal and nonverbal communication for maximum efficiency.

Ensuring emotional congruence between verbal and nonverbal communication is vital for maintaining the integrity of the interaction. A chat robot should, therefore, align its nonverbal expressions such as

tone, facial expressions, and body language with the emotional tenor of its spoken words to avoid confusing or misleading users.

## 4 Conclusion

This essay presents a framework for evaluating whether language models are truly communicating or simply mimicking human language. It introduces tasks designed to test models' understanding of intent, adaptability to context, and ability to maintain coherence over a conversation. The work also explores the design of chat robots that incorporate nonverbal cues like gestures and facial expressions to create a more natural and intuitive communication experience. The tasks focus on intention recognition, flexible application of knowledge across contexts, and context retention in extended dialogues. For embodied AI, the paper emphasizes the need for chat robots to use nonverbal communication effectively, aligning with verbal messages and enhancing information efficiency. This essay aims to advance AI towards more authentic and empathetic interactions with humans.

## References

- [1] Angeliki Lazaridou, Alexander Peysakhovich, and Marco Baroni. Multi-agent cooperation and the emergence of (natural) language. *arXiv preprint arXiv:1612.07182*, 2016. 1
- [2] Shuwen Qiu, Sirui Xie, Lifeng Fan, Tao Gao, Jungseock Joo, Song-Chun Zhu, and Yixin Zhu. Emergent graphical conventions in a visual communication game. *Advances in Neural Information Processing Systems*, 35:13119–13131, 2022. 2
- [3] Theodore R Sumers, Mark K Ho, Robert D Hawkins, and Thomas L Griffiths. Show or tell? exploring when (and why) teaching with language outperforms demonstration. *Cognition*, 232: 105326, 2023. 2
- [4] Michael Tomasello. *Origins of human communication*. MIT press, 2010. 1