Positioning AI-Mediated Communication in the Field of Human-Machine Communication

Hannah Mieczkowski
Department of Communication
Stanford University
Stanford, CA, USA
hnmiecz@stanford.edu

ABSTRACT

In the present paper, I outline relevant theories of social cognition, focusing on those that explain perceptions of others, and presiding ideas surrounding language production and comprehension in conversation, emphasizing linguistic alignment. Within these two broad components of communication, similarities and differences between human-machine communication (HMC) and AI-mediated communication (AI-MC) are investigated. Areas for future work in these research areas are also discussed.

CCS CONCEPTS

• Human-centered computing • Human computer interaction (HCI) • Interaction paradigms • Collaborative interaction

KEYWORDS

human-machine communication, AI-mediated communication, social cognition, linguistic alignment

ACM Reference format:

Hannah Mieczkowski. 2021. Positioning AI-Mediated Communication in the Field of Human-Machine Communication: Examples in Social Cognition and Language Use. In *Building Bridges and Not Walls: Expanding Human-Machine Communication Connections Within HRI, A Virtual Workshop at HRI 2021 (HRI'18). ACM, New York, NY, USA, 5 pages.* https://doi.org/XXXXXXXXXXXX

1 Introduction

The first models of communication, such as Shannon and Weaver's (1949) highlight one direction of information transmission: the sender to the receiver. Others, such as Dance's (1967) helical model, focus more attention on the dynamic nature of communication, emphasizing both the importance of interaction between people, as well how a conversation continues to change and grow as time passes. When Windahl and McQuail (1993) describe the helical model, they depict the conversation pair as "active" instead of "passive" participants in an interaction.

There have been fewer attempts to formally model the act or process of communication in the years that have followed. Instead, coupled with the expansive and interdisciplinary nature of the field of communication, scholars have continued to emphasize the role of social psychology, in addition to the role of language. Researchers of social cognition, or the "various psychological processes that enable individuals to take advantage of being part of a social group," have often focused on the impact of signals, which come from oneself, from others and from the messages they engage with in order to learn more about how people experience the world (Firth, 2008).

These phenomena are also apparent in the overlap between the studies of communication and linguistics, most notably through the concept of linguistic alignment in language production and comprehension. Linguistic alignment, or the state of "interlocutors hav[ing] the same representation" of language, at syntactic or semantic levels or both, is a core feature of communication (Pickering & Garrod, 2004). Without linguistic alignment, interlocutors could not even be considered as such - they would just be people engaging in simultaneous monologues. Linguistic alignment allows for the understanding others (e.g. what characteristics do I perceive in others if they are not aligned with me?), as well as the content of the conversation itself and what it references in the world.

In these models and research areas, humans are the focus, in that they are both the senders and receivers of the messages. Technology, whether it be the alphabet, a pen, or more modern forms of digital technology, is generally assumed to be the tool or medium through which people exchange messages. Overall, humans as communicators and technologies as tools has been the "dominant paradigm for communication research" (Guzman, 2018). However, as Guzman (2018) also notes, Shannon's (1948) original model of communication discussed the transmission of messages between humans and machines, not just between humans. Thus, one of the earliest and fundamentally influential models of communication discussed the reality of *human-machine communication* (HMC). Through Shannon's collaboration with Weaver, the original model was refitted to a social scientific

HRI'21, March, 2021, Virtual

https://doi.org/XXXXXXXXX

context instead of an engineering one (Rogers, 1997). As a result, subsequent definitions largely assumed that "communication must occur between two or more people" (Edwards et al., 2019), and machines were relegated to the role of a tool.

More recently, however, HMC has become a burgeoning subdiscipline in the larger field of communication, even though anthropocentric critiques have been common for several decades (e.g., "studying social relations without… nonhumans is impossible"; Latour, 1988). Similarly, Guzman (2018) argues that HMC is rightfully included in the field, as "the definition of 'communication' within human–machine communication and human–human communication is the same." By this, HMC scholars do not mean that interactions between humans and machines are "identical" to those between humans (Fortunati & Edwards, 2020), but rather the core aspects of communication - such as social cognition and linguistic alignment - are still applicable.

For the purpose of this paper, HMC refers to "meaning making [that] is a joint activity among human and machine agents," and most importantly, that "humans and machines are potentially equivalent interlocutors" who engage "in communication processes as both social and functional actors" (Banks & de Graff, 2020). This means that any human-machine interactions could include a machine, in any form, as one or more of the interlocutors, as long as one interlocutor is human.

However, because of technological advancements and increased machine autonomy, a machine acting as a "tool" is no longer necessarily a passive component of the communicative process. Unlike HMC, which emphasizes the role of a machine as a separate, distinct entity with which a human can converse, *AI-Mediated Communication* (AI-MC) reconceptualizes the role of machine as tool, highlighting how the medium itself can be an active participant in the conversation in the similar way that Windahl and McQuail (1993) would argue that people are.

Hancock, Naaman and Levy (2020) introduce AI-MC as "mediated communication between people in which a computational agent operates on behalf of a communicator by modifying, augmenting, or generating messages to accomplish communication or interpersonal goals." Whereas HMC only requires one human and one machine, AI-MC is concerned with machines that facilitate communication between at least two humans. Most importantly, the machine facilitating communication must be some form of artificial intelligence, or "computational systems that involve algorithms, machine learning methods, [or] natural language processing" (Hancock et al., 2020). Examples of this include suggested text responses like Google's smart replies, machine translation, voice assistants like Duplex that make calls on a person's behalf, AI-generated profiles on LinkedIn based on a person's employment background, or even deepfakes.

In addition to involving at least two people and an AI system, AI-MC must involve human control, and the AI system must be human-presenting. In the aforementioned examples, there is always at least one person "in control" of the AI system, meaning they are making decisions about how it will be used in interpersonal communication. Further, the AI system is ostensibly presenting itself as a human, not a machine. For instance, if Person A has an AI-generated profile on LinkedIn, the profile appears as if it is Person A and not something like a bot. The simplest form of AI-MC would include Person A using AI to present messages as themselves, but Person A could also use AI to present messages as if they were from another person (e.g. Person B).

Although many aspects of AI-MC overlap with and are relevant to the study of HMC more broadly, these two forms of communication raise distinct questions for their areas of study. In the present paper, I outline relevant theories in two research areas of communication: social cognition, focusing on those that explain perceptions of others, and language production and comprehension in conversation, emphasizing linguistic alignment. Within these two broad components of communication, similarities and differences between HMC and AI-MC are investigated.

2 Social Cognition

Although social cognition, by its human-focused definition, includes all psychological processes that underlie social interaction between people, interpersonal perception stands out as a uniquely and consistently appropriate topic that is applicable across communicatory contexts. Frameworks in face-to-face (FTF) communication, such as the stereotype content model (SCM) and behaviors from intergroup affect and stereotypes (BIAS) map have indicated that two crucial dimensions of interpersonal perception, warmth and competence, can predict both emotions and longer-term behavioral tendencies (Cuddy, Fiske & Glick, 2007).

Although in FTF contexts these perceptions are typically based on appearances, the constructs are applicable to technologically mediated situations as well (e.g. Walther, Loh & Granka, 2005). Regardless of the environment communicators are interacting in, Takayama (2009) notes that "perceptions are not second best to objective measures; in fact, it is perceptions and subjective realities that people judge and act upon." Beliefs about others, sometimes in spite of the true qualities of the others, are immensely powerful in social cognition.

Perhaps the most well-documented and generalizable interpersonal perception finding in HMC is the tendency for humans to treat their machine counterparts as if they were also humans. The Computers as Social Agents (CASA) framework's overarching prediction is that people will think about and act towards computers in much the same way that they think about and act towards other people (Reeves & Nass, 1996; Nass & Moon, 2000).

Although CASA specifically references computers, the principles underlying the framework have been applied widely to other types of machines. Research on CASA has replicated similar effects with conversational agents, robots, and chatbots (Eyssel & Hegel, 2012; Krämer et al., 2012). Much of this work has focused on human-robot interaction with social robots, likely due to their embodied nature; compared to text- or voice-based machines, robots can be designed to look "humanoid" with high degrees of anthropomorphism. As such, there are a number of similarities between human communication and human-robot communication when viewed through the lens of the CASA framework. For instance, perceptions of warmth and competence are made in much the same way for robots as they are for humans (Reeves et al., 2020). Further, many of the relationships described in the SCM and BIAS map regarding connections between impressions, emotions, and behavioral intentions are applicable to social robots (Mieczkowski et al., 2019).

Unlike HMC, social cognition in AI-MC involves at least three entities: two people and an AI system (a type of machine). As such, perceptions of "others" could involve both the other human communicator and the AI. In related work focusing on interactions between a person and a robot controlled by another person, participants in the study did not attribute personality traits or emotional states to the robot in question (Ling & Björling, 2020). This is in distinct contrast to the aforementioned work on impression formation of machines in HMC, where machines were perceived to have many qualities that humans also have.

However, perceptions of avatars, or virtual representations of people that are actually controlled by people, seem to conform to typical interpersonal perception expectations. People find anthropomorphic avatars more attractive and credible (Nowak & Rauh, 2005). In terms of behavior, a meta-analysis by Fox et al. (2015) indicates that people experience more social influence from avatars as opposed to agents, or virtual representations of people controlled by algorithms. Taken together, these studies suggest that there is not consensus among researchers regarding how machines operating on behalf of a person in communication between people might be perceived.

In addition to the impact of control on perceptions of machines, an important related construct is transparency of human or AI presentation. In a study of Airbnb profiles either written by a human or AI, people were least trusting of profiles labeled or presumed to be written by AI, but only when the profiles were among those also written by humans. Jakesch et al. (2019) termed this phenomenon the "replicant effect," but did not see this effect in instances where people engaged with all human- or all AI-composed profiles. Simply introducing an aspect of the unknown into the communication process impacted perceptions of trustworthiness substantially.

Although CASA suggests that increased humanlike qualities of a machine will prompt more natural, automatic social reactions - and AI controlled by a person could be considered very humanlike - the framework focuses on HMC, where the machine is an "other" in the interaction. The AI system in AI-MC may not necessarily be considered an "other," and as such, may invoke perceptions different from that of a machine in typical HMC.

3 Language Use

Social cognition and language use are inextricably linked. Just as people form impressions of each other and themselves based on cues such as appearance, vocal features, and gender, people use language as a way to make judgments and decisions in human communication. Linguistic alignment in particular necessitates communication between at least two people, as it is a measure of how much conversation participants "converge… on common ways of speaking" (Pickering & Garrod, 2004). Linguistic alignment is an automatic and dynamic process that occurs over the time frame of a conversation and can carry over into future conversations. Alignment in conversations happens on multiple levels, such as in accents and speech patterns (Cappella & Planalp, 1981), syntax (Levelt & Kelter, 1982), as well as content (Brennan & Clark, 1996).

An important facet of linguistic alignment is that it generally relies on communicators cooperating with each other, some openness to the content that the other participant(s) are discussing, and adherence to social norms. As indicated by Breazeal (1999), "humans expect to share control with those whom they socially interact" and shared control is only possible through "empathetic understanding of others' mental states" (Kozima, Nakagawa & Yano, 2004).

Mental states and minds are typically inferred by human communicators when they interact with machines or AI, and largely automatic processes like linguistic alignment are less likely to be influenced by the mental state of a conversation partner (Branigan et al., 2003). As such, we might expect conversational dynamics in HMC and AI-MC that are similar to those in human communication, with potential differences rooted in beliefs about the mental capabilities of the involved machines.

Overall, past research largely indicates that people linguistically align, both verbally and nonverbally, with machines in much the same way that they do with people. These comparable findings across communication settings and levels of language production and comprehension contribute to the aforementioned automatic model of linguistic alignment, which argues that linguistic alignment is automatically triggered through processes such as residual activation or implicit learning (Chang 2006; Pickering & Branigan, 1998). However, the way humans perceive machines may be the root of the differences that do exist in linguistic alignment between the two communication types. Substantial research suggests that humans not only linguistically align with machines, but they linguistically align *more* with machines than they do with other people. Branigan et al. (2010) claim that this pattern is due to peoples' "beliefs about the... linguistic capability of their interlocutor" and since "humans are generally believed to be more capable… in the domain of communication," people attempt to match the communicative abilities of their mechanistic conversation partner.

The process of linguistic alignment has the potential to become slightly more complex in AI-MC. If linguistic alignment is converging on common ways of speaking - including semantically, syntactically, nonverbally and otherwise - then the AI system modifying or generating communication could interrupt or magnify this process. Especially if a communicator has minimal digital literacy, the introduction of an AI mediator into a conversation could provoke more thoughtful, as opposed to instinctive, responses. In this way, AI-MC could be distinct from both human communication and HMC because of the decreased automaticity of the linguistic alignment process.

However, even a confusing addition to the communication process may not be enough to override the influence of linguistic alignment. As previously noted, machines introducing new terms into conversations actually strengthened alignment, likely due to perceptions of the machine's capabilities. Recent work by Hohenstein et al. (under review) indicates that when the text of AI-generated suggested text responses was manipulated to be either positive or negative, the sentiment of the conversation as a whole became more positive or negative, respectively. Further, some preliminary analyses of text data from Mieczkowski et al. (unpublished) suggest that linguistic alignment occurs to approximately the same extent in conditions where there was no AI system involved and when the conversation was AI-mediated.

Taken together, these findings provide indications that linguistic alignment may occur both between people and between the communicators and the AI system. One way to investigate AI or human influences on linguistic alignment in AI-MC would be to manipulate the language of one human conversation partner, as well as the AI system. For instance, a human confederate could use more positive language, and an AI system could suggest more negative language, and changes in the participant's message sentiment over the course of a conversation could be observed.

Current research on AI-MC has focused almost exclusively on sentiment in text-based AI systems due to their prevalence and increasing integration into daily communication. However, AI-generated speech and visuals, in both benign and malignant settings will be more widely available in the near future. Investigating differences in linguistic alignment based on technological affordances and, relatedly, perceptions of humanness, will be necessary to properly grasp the boundary conditions of AI-MC's impact on language use.

4 Conclusion

Understanding how key components of the human communication processes, such as social cognition and language use, can be applied to HMC and AI-MC provide extensions and boundaries for theories of communication. Consistent empirical findings in human interpersonal perception are often easily applied to HMC. Using CASA as a framework, the replications of human communication processes in human-machine interaction can be explained by the tendency for people to perceive and react towards social agents in much the same way that they do to people, and this effect is magnified by the humanlike qualities of a social agent. Interpersonal perception in AI-MC has a similar basis in human communication theories and frameworks but differs in the minimum number of involved parties (human and machine as opposed to at least two people and an AI system). Additionally, AI systems are controlled by a human, and necessarily perceived to be humans by one of the parties, potentially providing further complications for impression formation.

Theories of interactive linguistic alignment are also applicable to HMC and potentially AI-MC, with a few notable caveats. Although machines and AI systems may be perceived as human, people often linguistically align with these technologies in a way that indicates the limits of their language capabilities. As such, in HMC, alignment follows patterns similar to those in human communication where one conversation participant is not a native speaker of the relevant language. This area is underexplored in AI-MC at present, but technological capabilities and perceptions of humanness are likely to be relevant. Future work in this area might parse apart the influences of the AI system and the conversation participant in communication over time.

REFERENCES

[1] Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication* (12th Printing). Urbana, IL: University of Illinois Press.
[2] Laswell, H. (1949). "The Structure and Function of Communication in Society", Mass Communication, ed. Schramm, Wilbur. *Urbana, University of Illinois Press*, 102-115.
[3] Dance, F. E. (1967). A helical model of communication. *Human Communication Theory, New York, NY: Holt, Rinehart and Winston*.
[4] Windahl, S., & McQuail, D. (1993). *Communication models for the study of mass communications*. Longman.
[5] Frith, C. D. (2008). Social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1499), 2033-2039.
[6] Pickering, M., & Garrod, S. (2004). The interactive-alignment model: Developments and refinements. *Behavioral and Brain Sciences*, 27(2), 212-225. doi:10.1017/S0140525X04450055
[7] Guzman, A. L. (2018). What is human-machine communication, anyway. *Human-machine communication: Rethinking communication, technology, and ourselves*, 1-28.
[8] Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3,4), 379–423, 623–656.
[9]Rogers, E. M. (1997). *A history of communication study: A biographical approach*. New York, NY: Free Press.
[10]Edwards, A., Westerman, D., Edwards, C., & Spence, P. R. (2019). Communication is… Transhuman. In A. Tyma and A. Edwards (Eds.) *Communication Is… Perspectives on Theory* (pp. 47–66). Cognella.
[11] Latour, B. (1988). How to write The Prince for machines as well as for machinations. *Technology and social change*, 20-43.
[12] Fortunati, L., & Edwards, A. P. (2020). Opening space for theoretical, methodological, and empirical issues in human-machine communication. *Human-Machine Communication*, *1*(1), 1-18.

[13] Banks, J., & de Graaf, M. (2020). Toward an Agent-Agnostic Transmission Model: Synthesizing Anthropocentric and Technocentric Paradigms in Communication. *Human-Machine Communication*, *1*(1), 19-36.
[14] Hancock, J. T., Naaman, M., & Levy, K. (2020). AI-Mediated Communication: Definition, Research Agenda, and Ethical Considerations. *Journal of Computer-Mediated Communication*, *25*(1), 89-100.
[15] Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect and stereotypes. *Journal of personality and social psychology*, *92*(4), 631.
[16] Walther, J. B., Loh, T., & Granka, L. (2005). Let me count the ways: The interchange of verbal and nonverbal cues in computer-mediated and face-to-face affinity. *Journal of language and social psychology*, *24*(1), 36-65.
[17] Takayama, L. (2009, March). Making sense of agentic objects and teleoperation: In-the-moment and reflective perspectives. In *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 239-240). IEEE.
[18] Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge university press.
[19] Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of social issues*, *56*(1), 81-103.
[20] Eyssel, F., & Hegel, F. (2012). (S)he's got the look: Gender stereotyping of robots. *Journal of Applied Social Psychology*, 42(9), 2213–2230.
[21] Krämer, N. C., von der Pütten, A., & Eimler, S. (2012). Human-agent and human-robot interaction theory: Similarities to and differences from human-human interaction. In M. Zacarias & J. V. de Oliveira (Eds.), *Human-computer interaction: The agency perspective* (pp. 215–240). Berlin, Germany: Springer-Verlag.
[22] Reeves, B., Hancock, J. & Liu S.X. (2020). Social robots are like real people: First impressions, attributes, and stereotyping of social robots. *Issue 1*, *1*(1).
[23] Mieczkowski, H., Liu, S. X., Hancock, J., & Reeves, B. (2019, March). Helping not hurting: applying the stereotype content model and bias map to social robotics. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 222-229). IEEE.
[24] Ling, H. Y., & Björling, E. A. (2020). Sharing stress with a robot: what would a robot say?. *Human-Machine Communication*, 1(1), 8.
[25] Nowak, K. L., & Rauh, C. (2005). The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction. *Journal of Computer-Mediated Communication*, *11*(1), 153-178.
[26] Fox, J., Ahn, S. J., Janssen, J. H., Yeykelis, L., Segovia, K. Y., & Bailenson, J. N. (2015). Avatars versus agents: a meta-analysis quantifying the effect of agency on social influence. *Human–Computer Interaction*, *30*(5), 401-432.
[27] Jakesch, M., French, M., Ma, X., Hancock, J. T., & Naaman, M. (2019, May). AI-mediated communication: How the perception that profile text was written by AI affects trustworthiness. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).
[28] Cappella, J. N., & Planalp, S. (1981). Talk and silence sequences in informal conversations III: Interspeaker influence. *Human Communication Research*, *7*(2), 117-132.
[29] Levelt, W. J., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive psychology*, *14*(1), 78-106.
[30] Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482.
[31] Breazeal, C. (1999). Robot in society: Friend or appliance. In *Proceedings of the Agents99 workshop on emotion-based agent architectures*.
[32] Kozima, H., Nakagawa, C., & Yano, H. (2004). Can a robot empathize with people?. *Artificial Life and Robotics*, *8*(1), 83-88.
[33] Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F., & Nass, C. (2003, July). Syntactic alignment between computers and people: The role of belief about mental states. In *Proceedings of the 25th annual conference of the cognitive science society* (pp. 186-191). Lawrence Erlbaum Associates.
[34] Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological review*, *113*(2), 234.
[35] Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and language*, *39*(4), 633-651.
[36] Branigan, H. P., Pickering, M. J., Pearson, J., & McLean, J. F. (2010). Linguistic alignment between people and computers. *Journal of pragmatics*, *42*(9), 2355-2368.
[37] Hohenstein, J., DiFranzo, D., Kizilcec, R., Aghajari, Z., Mieczkowski, H., Levy, K., Naaman, M., Hancock, J., Jung, M. (under review). Artificial intelligence in communication impacts language and social relationships.