Journal of Media Horizons
Volume 6 Issue 1, 2025 ISSN: 2710-4060 | 2710-4052

# PRAGMATICS OF MULTIMODAL COMMUNICATION: ASSIMILATING LANGUAGE OF GESTURES AND GAZE

**Dr. Rai Zahoor Ahmed[*1] Ayesha Saleem[2]**

*[*1]Assistant Professor, COMSATS University Islamabad, Lahore campus*
*[2]MS English Linguistics Scholar, COMSATS University Islamabad, Lahore campus*

[*1]rai.zahoor@cuilahore.edu.pk, [2]fa24-mel-010@cuilahore.edu.pk

Corresponding Author: *

**Received:** November 21, 2024 **Revised:** December 21, 2024 **Accepted:** January 06, 2025 **Published:** January 13, 2025

**ABSTRACT**

*Pragmatics plays a vital role in human speech communication in real life situations. The multifunctionality of multimodal communication is integrated with verbal language and other non-verbal communication through gestures and gaze, which enhances linguistics understanding in several contexts. Gestures remove mental overload, enhance memory and make the notion easier to understand. Gaze can also be used to control flow of conversation and create social contact among human beings. All these modes blend synergistically in an effort to create a communication system that is better than the sum of parts. Focus and base on the neural foundations of multimodal communication by its roots in evolution along with applications in school, technology, and professionalism, and with virtual and augmented reality emergence creating further importance towards multimodal systems. However, cross-culture differences and multimodal integration have remained and are rising in many different contexts as problems. This study opens up such possibilities for an extended share of social and cognitive benefits. The research is particularly valuable for developing educational tools of innovative technological and inclusive communication methods adapted to diverse, multilingual target audiences.*

***Keywords:*** *Multimodal communication, Non-verbal cues, Cognitive processing, Gesture, Gaze, Cross-cultural differences.*

## INTRODUCTION

Beyond the simple exchange of words, communication is a multifaceted phenomenon. Humans use a variety of non-verbal clues, including gaze and gestures to enhance and support the spoken language to make it more meaningful. Multimodal communication is the term for the mix of spoken and non-spoken components that enables more complex and efficient information transfer (Ahmed &Ashraf, 2024). While gaze helps to control social interaction by indicating attention, emotion or conversational terms, gestures – which can range from basic hand movements to intricate mimetic actions, often reinforce or clarify spoken information (Kendon,2020; Kleike 2022). When combined these modalities improve communication and comprehension both in social and professional context.

Recent studies in communication technology, psychology and cognitive science have overall contributed to a growing interest in the "study of multimodal communications". It has been demonstrated that gestures are essential to cognitive processing, helping listeners and presenters comprehend and retain information (Alibali & Heath, 2021). According to LaFrance and Hecht (2023), gaze is also a crucial for directing social contact, indicating when

someone is taking turn in a discussion, and controlling the flow of speech.

Human communication has been built on the co-occurrence of verbal and non-verbal messages. Though text is more structured by the linguistic system, face-to-face communication involves several sensory modalities simultaneously. Speech is complemented by the gestures and gaze of the speaker that can make information exchange richer and interactive. For instance, disambiguating complex verbal statements through the gestures or attention through gaze are some signals of a lively and more dynamic communication.

Multimodal communication also functions as a crossing from abstract to concrete. The verbal language is very abstract, but the body's gestures give it concrete visual and spatial meanings of the ideas (McNeill, 1992). Gazing also controls attention and encourages social involvement by adding some rhythm to the conversation and when it is time to stop and listen or speak (LaFrance & Hecht, 2023). It therefore comes out as a very adaptable, resilient, and fit system of communication for a wide variety of contexts and audiences.

Gestures are used universally across cultures and contexts and carry functions ranging from focusing speech to replacing it at times. The research shows that there are three basic types of gestures in interaction with speech, namely deictic gestures, which involve pointing to objects or directions; iconic gestures, which represent the semantic content of speech; and beat gestures, which are the rhythmic movements that emphasize the prosody of speech.

Iconic gestures are specifically effective to enhance comprehension in a more nuanced or ambiguous discussion. For instance, if someone were teaching a classroom what a pendulum is, she could show exactly what that design does-swing-with her hand. Such gestures clarify speech and thus serve as an additional device for memory for the receiver. According to research by Alibali and Heath (2021), gestures actually facilitate cognitive processing by reducing the listener's cognitive load, thus giving him more time to integrate the verbal and visual inputs more effectively.

In addition to their cognitive uses, gestures allow for incorporation in communication. They create a connection for the language-disabled or in multilingual contexts where the verbal language is not common. Through gestures, therefore, it ensures that more people have seen the message through signs.

While gestures are visual aids for speaking, the role of glance is very crucial in regulating the social interaction. Eye contact is a pointer to attention, emotion, and social intent, hence a non-verbal cue for interpersonal dynamics. It helps in establishing a relationship between the speakers and listeners, therefore aiding mutual understanding and collaborating on conversations. This may also be used in managing turn-taking, indicating that one should speak or if the speaker has finished making his point (LaFrance & Hecht, 2023).

This term refers to the form of communication where the verbal elements integrate with the non-verbal information like gazing, gestures, or face expressions in order to provide meanings for the audience effectively. In this case, in particular iconic gestures are vital and depict content in concrete visual forms of that speech given (McNeill, 1992).

Multimodal communication means bringing in more significant meaning with the use of the verbal mode and non-verbal modes of communication like gestures, facial expressions, gaze, and prosody. In that respect, it gives critical roles for the improvement of interaction, which enables easy understanding and keeping better memories (Goldin-Meadow, 2003). Such method represents the complexity in human communication since it applies to several sensory inputs in reaching a richer and dynamic way of information interchange.

Gaze is notably vital in joint attention - some sort of shared attention that exists over some object or concept making the heart of effective communication. It is mostly needed for parent-child relationships, in which an infant counts on his or her parent's gaze to understand this world and a language; for professional talk, it acts as a cue showing interest, agreement, or disagreement which shapes the way of discussion and its results.

Apart from the social functions, gaze supports cognitive processes. It has been proven that when the speaker maintains eye contact, then there is a greater likelihood that the information would be remembered by the listener. Gaze works in

concert with gestures to point attention towards the listener. A speaker might look at an object and point at it so that the listener's attention would be drawn to the same point of reference.

Multimodal communication is the tool of cognitive processing. For instance, gestures lighten the cognitive load of an individual by externalizing what is going on in one's head so that the brain can give more resources to complex problem-solving (McNeill, 1992). Eye contact and facial expressions are the non-verbal cues for attention and emotional intent in social contexts that foster rapport and cooperation among individuals (Argyle, 2010). The integration of gestures, gaze, and speech constitutes a multimodal system greater than the sum of its parts. Such modalities work together dynamically to compensate for each other's weaknesses. For example, gestures may clarify ambiguous verbal statements, while gaze will ensure that the listener's attention is concentrated on the right parts of the message. This synergy enhances not only comprehension but also the emotional and social resonance of communication. This integrates gestures and gaze with talk in business settings such as business presentations or negotiations to elicit better clarity and forcefulness in the message or communication. For example, a presenter using gestures toward key points with eye contacts to their audience is engaging and has more of leaving an impression to the participants. The same modalities are fostering empathy and association in social contexts, letting people get along well in any emotional or social landscape to be explored.

This growing interest in multimodal communication can be noted in the most recent developments of psychology, communication technology, and cognitive science. Evidence suggests that involving gestures and gaze in a conversation enhances both the cognitive and social advantages of a discourse. In this aspect, better users of gestures and gaze by speakers tend to keep their audience engaged and maintain their attention better. The listeners also obtain higher understanding and memory retention (Alibali & Heath, 2021).

As discussed above, technology has increasingly adopted the concept of multi-modal communication. The advanced realization of VR and AI enables natural and intuitive interfaces with growing prominence of gesture and gaze recognitions. For example, through virtual agents imitating human-like gestures, one develops familiarity and then trust in certain practices, like education, therapy, or customer services.

Despite the great strides in research, multimodal communication remains an emerging field of study. There is continued interest in understanding how different modalities interact with each other in diverse cultural, linguistic, and technological environments. For example, cross-cultural differences in the use of gaze and gestures underscore the role of context in designing a communication strategy or technology. Also, knowledge of the brain processing these modalities can lead to insights into how human beings acquire and fine-tune communication.

Combining these modalities with spoken language improves social cognitive results and facilitates more efficient communication. The goal of this research paper is to present of how language, gesture, and gaze can be integrated into multimodal communication. Researcher will examine how these components work together to promote social interactions and assist cognitive processes by going over current findings. This literature review will describe the state of the art findings on how these elements combine in order to further the possibility of social interaction with better cognitive processes and outcomes of communication. Such a holistic approach makes progress in understanding human communication and provides pragmatic insight that can be applied in educational and technological environments, as well as clinical practice.

**Research Objectives**

1. To explore the integration of verbal language with gestures and gaze in communication.
2. To assess the cognitive benefits of multimodal communication

**Research Questions**

1. How do gestures complement and enhance the meaning of spoken language in communication?
2. What cognitive benefits arise from the integration of language, gestures and gaze in communication?

**Literature Review**

Multimodal communication includes various mode combinations such as speech, gestures, facial expressions, and body language in interaction. This reflects the complexity of human interaction and its cognitive, evolutionary, and practical implications.

Multimodal communication stresses the integration of different modalities into a unified system. According to McNeill's growth point theory, gesture and speech stem from a single conceptual system where linguistic and imagistic information dynamically interact (Clough & Duff, 2020). Similarly the sketch maintains that although speech and gesture share the same communicative intention, they are processed in parallel systems that cooperate to deliver a coherent message. This integration depicts that language is not verbal alone but is multimodal, involving auditory, visual and motor formation (Townsend et al, 2019).

Multimodality casts light on the origin of human communication. Perhaps language originated as a multimodal system in which gesture and vocalization were parts of the system (Fröhlich et al., 2019) .This view bridges the traditional debates over whether gestures or vocalizations were the first communication systems. For example, whereas gestures play informative role in close range communication among non-human primates, vocalization dominates in humans indicating a shift in "carrying roles" over evolutionary time.

Gestures supplement and even enrich speech with spatial, imagistic, and abstract meaning. For example, iconic gestures represent the semantic content of speech. This even further enhances comprehension, particularly in noisy environments (Krason, 2022). Gestures also have cognitive roles, such as memory facilitation, problem-solving and language learning. Further researches on patients with aphasia have shown that gestures may complement the deficits in speech production. This implies a rehabilitative use. Similarly visual input which involves facial expression and gaze, is very important for interpreting speech. The McGurk effect shows how the auditory impression changes through visual input, hence showing dependence of modalities on each other (Townsend et al., 2019). Krason (2022) further demonstrates that listeners prefer visual speech cues, especially in challenging auditory conditions.

According to Krason (2022), gestures and mouth movements play different yet complementary roles in speech interpretation. While the former is very useful for building situational models of discourse, the information to be derived from mouth movement cues is phonological yet useful for noisy conditions. It would then be observed that multimodal communication is dynamic and interactive in that different cues might have more priority than others, depending on the need or situation.

Integration of multimodal inputs occurs in complex networks. The superior temporal sulcus is also a convergence point for audio and visual signals as it is directly in communication with both of them. Recent studies have shown that simultaneous processing of speech with gestures renders cognitive representation resilient and hence, easy to process language (Clough & Duff, 2020). This has practical applications in numerous fields. For example in education, multimodal approaches enhance learning, particularly for bilingual students or hearing-impaired individuals (Krason, 2022). Similarly, for human computer interaction, there is added improvement in experience through designing gesture and speech input systems while being integrated in design models that stimulate human multimodal behaviors (Cassel & Prevost, 1994).

Traditional linguistic frameworks that view language as a speech-centric, unimodal system are challenged by multimodal communication. Rather, it suggests that a synthesis of modalities leads to communication, resulting in a more varied and flexible information flow. According to McNeill (1992), the growth point theory posits that gestures and speech are both derived from the same conceptual system. Speech provides symbolic, linear representations, while gestures provide holistic, spatially organized representations, which are tightly integrated to enhance the intelligibility of the message (Clough & Duff, 2022). For example, iconic gestures can represent the action or object being described by speech, so the recipient of the message can also simultaneously process abstract and visual information (Krason, 2022).

In sketch modal, though speech and gestures are differentiated yet interactive routes that serve the

same communicative purpose. According to Clough and Duff, this approach places on emphasis on the interaction of these two modalities with respect to their complementary functions in filling up each other's inadequacies for individuals suffering from impairments (2020). This serves as an affirmation for both of these theories when emphasizing the multimodal basis of communication with each contributing channel reinforcing the others to develop a coherent stream of meaning.

Multimodal communication sheds light on the evolutionary roots of language. Comparative studies have shown that humans are nonhuman primates have a common ancestral origin through the display of multimodal signaling behaviors (Fröhlich et al., 2019). For instance, in primates, gestures and vocalizations are used together and share properties important to language, such as intentionality, reference and turn-taking (Townsend et al., 2019). While gestures lead close-range communication among primates, the role of carrying information from these animals has been taken by vocalization in the process of evolution. This means there was increased adaptive fitness because the obstruction of visual signals for long- distance communication (Fröhlich et al., 2019).

It reconciles the "gesture-first" theory of language origins, suggesting that manual gestures were the modal probably primarily engaged in early communication, with vocal theories as it views that language probably evaluates as an integrated system with gestures and vocalizations each complementing the order for the maximum in communicative efficiency. According to Townsend et al, (2019) gestures are an integral part of multimodal communication, bringing unique advantages to meaning conveyed. Iconic gestures, which graphically depict the content of speech, are very effective in illuminating abstract or complex ideas. For example, pointing with two fingers while saying "She went to the park" helps explain by visualizing the activity (Cassell & Prevost 1994). Besides being an active component of communication, gestures also have an active role in cognition. According to studies, gestures enhance memory and assist in solving problems as well as in language acquisition among children (Clough & Duff, 2020). Furthermore, the information given by "unspoken" words opens a speaker's thoughts (McNeill, 1992). For example, if gestures and speech deliver contradictory or complementary information with each other, then what the listener receives can be synthesized into a coherent mental representation (Krason, 2022).

Visual speech cues involve mouth movements, eye gaze, and facial expressions. It is in fact established through research that such cues play a more significant role in adverse listening environments where degraded auditory signals are present. For instance, the McGurk effect is one of the examples of deep integration of audio-visual information in the brain, where visual inputs modify auditory perception (Townsend et al., 2019).

Krason's research, 2022 on how mouth is informative found the role of visual cues working in dynamic functions in term of context. Although gestures typically augment the discourse, their significance is more critical only under noisy conditions where speech phonological cues are perceived with the help of mouthing movements. Flexibility provides insights into how communication using different modes accommodates different situation. The neural foundation of multimodal communication is seen in dependence on distributed networks of the brain that take in and integrate sensory, motor and linguistic inputs. One important node for combining visual and auditory signals is the superior temporal sulcus, an area critical for the processing of both speech and gestures together (Clough & Duff, 2020). Neuroimaging experiments have further demonstrated that processing gestures activates areas that are associated with motor preparation, making language an embodied phenomenon after all (Townsend et al., 2019).

Embodied cognition is another theory that extends to explaining how language processing has roots in sensory and motor experience. For example, when a person reads words with an action-related meaning, there will be corresponding activations of motor areas in the brain linking linguistic representations with actual actions. This neural overlap is what underscores the intermodal construction of meaning (McNeill, 1992; Krason, 2022).

From evolutionary grounds, human communication has been based on the argument that the multimodal signaling system of primates involves gestures and vocalizations in signaling intent (Tomasello, 2008). From the neural perspective, it has been proved that superior temporal sulcus plays a key role in combining the auditory and visual information with a biological basis for the multimodal communication system (Hickok & Small, 2015).

In human-computer interaction, multimodal systems like those that rely on gesture recognition and voice commands make the user experience better, more intuitive interfaces, by (Cassell, 2001). In an educational environment, a multimodal approach like visual aids and interactive gestures helps learners with hearing impairments, especially language learners, (Mayer, 2009).

Multimodal communication is not a characteristic of human language but rather more vital to enrich communication, understanding, and learning. Complexity in human interaction within this context refers to the scope of multimodal communication from its evolutionary origins up to neural underpinnings and practical applications. Other promising study roles within diverse contexts may include clinical settings and technological interfaces.

Despite the above significant advancements, challenges in the study of multimodal communication persist. First, more research is still required to determine how best different modalities interact and complement each other in multicultural and multilingual settings. Additionally, neural mechanisms involved in integrating the different modes of communication remain unclear and should be elucidated; particularly in individuals with communication impairments. Consequently, the need for further research that addresses how the potential benefits of multimodal communication in modern virtual and augmented realities may be enhanced.

**Research Methodology**

Qualitative content analysis was employed to the speech of Imran Khan in the 74th United Nations General Assembly (UNGA) on September 27, 2019. The method used is suitable for answering the research questions in that it encompasses both linguistic and multimodal aspects of the speech - verbal language, gestures, facial expressions, and intonation. The whole impact and underlying themes behind this speech would be understood more holistically.

**Data Collection**

The main video recording of Imran Khan's speech at the UNGA has been used as the primary source for this research. The speech has been sourced from freely accessible archives to ensure transparent and authentic data. The same has been transcribed in detail to enable textual analysis and coding. In addition, gestures, gaze, and body language of the speaker have been documented based on a multimodal communication framework. Multimodal Discourse Analysis, this method was applied to analyze how gestures, tone, and other non-verbal aspects complement and augment verbal communication. For instance, it placed emphasis on whether the spoken words were congruent or divergent from accompanying gestures or expressions. Techniques such as McNeill's growth point theory were applied to explain how gestures augment the semantic content of speech .

**McNeill's application growth point theory**

According to McNeill's Growth Point Theory, speech and gesture have a common origin from one source of concept. Gestures are not additions but are parts that are integrated into the process of creating general meaning. Therefore, deduce how Imran Khan's gestures supplemented the semantic values of his verbal messages.

**1. Gestures Confirm Urgency in Climate Change**

**Content speech:** Regarding climate change, Khan highlighted its urgency and scale of the crisis it is posing to the vulnerable countries like Pakistan.

**Gesture Analysis:** He spoke using sweeping motions with his hand and very dramatic pointing of the arm when mentioning global warming and melting glaciers. This was to signal a dimension of the problem, visually illustrating the extent.

**Growth Point:** The conceptual growth point here is in the simultaneity of verbal urgency ("time is running out") and the wide-spanning gestures that conveyed an expansive, global threat.

**2. Religious Discrimination Illustration**

**Content of Speech:** Khan talked about Islamophobia, elaborating how Muslims are alienated and discriminated upon.

**Gesture Analysis:** As he described marginalization, he enacted enclosing hand gestures, isolating himself. Sometimes he extended his hands out of the enclosure, signifying breaking the barriers, hence matching his call for understanding and tolerance.

**Growth Point:** The hand gestures demonstrate the contradiction between oppression and freedom. Thus, a visual representation of the difference between the problem and the proposed solution.

**3. Gestures Supporting the Kashmir Conflict Narrative**

**Speech Content:** The Kashmir issue framed it as a human rights crisis and a potential nuclear flashpoint.

**Gesture Analysis:** Criticizing the Indian government's actions, Khan used sharp, downward slicing motions, which visually mimed the severity of his words. While appealing for UN intervention, open-palmed gestures signified a plea for justice.

**Growth Point:** The chopping and slashing gestures symbolized how people of the place felt regarding losing liberties in Kashmir, while the open-palmed gesture complemented the oral plea to make a harmonic criticism and appeal.

**4. Economic Inequality and Exploitation of Wealth**

**Speech Content:** She spoke against tax havens, which create leakage of wealth from the Third World countries.

**Gesture Analysis:** The hand patterns and circular motions indicate rounds of exploitation, while upturn gestures point to aspirations toward reform for equitable distribution.

**Growth Point:** The hand gestures used to illustrate cycles and transformations were just perfect in his speech to depict the dynamic nature of economic systems.

**5. Acts of Leadership Responsibility**

**Speech Content:** Khan has asked world leaders to act morally and decisively.

**Gesture Analysis:** This is the time of pointed gestures toward the audience. Thus, he is giving them metaphoric call-to-action because the moments of silence add up to the impact of his saying.

**Growth Point:** The gestures are pointed as an extension of his verbal appeals, which is the directive nature of his message.

Applying McNeill's Growth Point Theory, it may be viewed that Imran Khan's gestures are the integral part of his speech. These not only develop the semantic content but also provide the multimodal aspect that will support his rhetoric goals. His gestures would be in the form of visual language, connecting the abstract notions with the action-oriented outcomes and intensifying the delivery of the message to the audience.

**Findings**

**1. Integration of Verbal and Non-Verbal Cues:**

Multimodal communication combines verbal language with gestures and gaze, thereby greatly enriching the flow of information.

Gestures help explain ambiguous statements and promote better memory recall by lightening cognitive load. They can be visual aids for the verbal content, as with iconic gestures that depict abstract or complex concepts.

**2. Social and Cognitive Functions of Gaze:**

Gaze governs the flow of a conversation and creates social ties by showing attention and emotional intent. Joint attention of multimodal communication through gaze facilitates successful communication that aids in establishing a mutual understanding of contexts at the professional and individual level.

**3. Multimodal Gestures:**

There are three different categories of gesturing-deictic, iconic, and beat serving distinct purposes including pointing to enable prosody and enhance intelligibility. The iconic gesture represents visual as well as spatial visualization; this has made it

play a major cognitive processing role, simplifying abstract thinking.

**4. Multimodal Communication Professional and Social Context:**

The integration of gestures and gaze in speeches or presentations can make the audience more involved and have better retention. These modalities help foster empathy and association, which will guide speakers through various emotional or social landscapes.

**5. Evolutionary and Neural Considerations:**

Human communication must be a multimodal system of gesture and vocalization to best serve its purpose. Neural studies indicate that areas, such as the superior temporal sulcus, are multimodal processing sites, thereby setting a biological basis for an integration of speech, gestures, and gaze.

**6. Utility and Challenges:**

Multimodal approaches improve educational aids, particularly for bilingual speakers or hearing-impaired learners, and improve human interaction with computers. Cross cultural differences in gestures and in gaze necessitate context- sensitive communication strategies. Some of the remaining work is research on how these modalities can be combined in multilingual and virtual contexts.

**Conclusion**

The combination of gestures, gaze, and speech in communication creates a synergistic multimodal system, which is more than the sum of unimodal approaches. It includes verbal and non-verbal cues in multimodal communication, which improves comprehension, develops social bonding, and assists cognitive processes. This relates to very diverse fields such as education, technology, clinical settings, and professional interactions.

Gestures and gaze complement the verbal language but play distinct roles: gestures make complex ideas easier to see, while gaze establishes relational dynamics and attention. This double role supports an inclusive communication framework that is particularly beneficial for language-impaired people or people working in multilingual settings.The creation of human-like experiences in interaction through multimodal communication and recent advancement into virtual and augmented realities gives more importance to communication in the context of the modern era as well as future orientation. Despite the progress made, the knowledge areas suffer gaps; in knowing how multimodal cues work across cultural and linguistic boundaries and in exploring neural underpinnings. Such research gaps should be addressed in subsequent studies to optimize use of multimodal communication for social and cognitive benefits across multiple settings. The studies establish the potential of a dynamic and evolving field toward advancement of human interaction and technological innovation.

**REFERENCES**

Ahmed. Z., Ashraf. S.(2024). Role of Pragmatics in Cross-Cultural Communication: A Focus on Politeness. Policy Research Journal, 2(4), 2231-2239.

Alibali, M. W., Heath, C. (2021). The cognitive functions of gestures: Insights from psychology and communication. Cognitive Science Review, 45(3), 342–357.

Cassell, J., Prevost, S. (1994). Embodied communication in human-computer interaction: The role of gesture and multimodal systems. Journal of Interactive Communication, 23(2), 123–138.

Clough, M., & Duff, P. (2020). Multimodal communication: Interactions between gesture, speech, and the brain. Journal of Cognitive Science, 15(4), 78–91.

Fröhlich, M., Townsend, S. W., & Zuberbühler, K. (2019). Evolutionary perspectives on multimodal communication: Linking gestures, vocalizations, and cognitive capacities. Trends in Neuroscience, 42(1), 13–25.

Kendon, A. (2020). Gesture: Visible action as utterance. Cambridge University Press.

Kleike, L. (2022). Gaze and turn-taking in social interactions: A cross-cultural analysis. Journal of Human Interaction Studies, 14(1), 56–72.

Krason, A. (2022). The role of visual cues in speech interpretation: A multimodal perspective. International Journal of Communication Research, 16(3), 89–105.

LaFrance, M., & Hecht, M. A. (2023). Eye contact as a social and cognitive signal:

Insights from cross-disciplinary studies. Communication Research Quarterly, 35(2), 120–134.

McNeill, D. (1992). Hand and mind: What gestures reveal about thought. University of Chicago Press.

Townsend, S. W., & Zuberbühler, K. (2019). Multimodal origins of human communication: A comparative analysis. Journal of Language Evolution, 4(2), 110–128.

Argyle, M. (2010). Bodily communication. Routledge.

Cassell, J. (2001). Embodied conversational agents: Representation and intelligence in user interfaces. AI Magazine, 22(4), 67–84.

Goldin-Meadow, S. (2003). Hearing gesture: How our hands help us think. Harvard University Press.

Hickok, G., & Small, S. L. (2015). The neurobiology of language. Academic Press.

Mayer, R. E. (2009). Multimedia learning. Cambridge University Press.

McNeill, D. (1992). Hand and mind: What gestures reveal about thought. University of Chicago Press.

Tomasello, M. (2008). Origins of human communication. MIT Press.