ChinaRxiv 中X

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-201911.00004

A New Perspective on the Cognitive Functions of Gestures: The 'Spatialization'Gesture Hypothesis

Authors: Yu Wenhua, Lu Zhongyi, Lu Zhongyi

Date: 2019-11-04T00:00:00+00:00

Abstract

Humans often accompany speech or thought with gestures. Gestures are spontaneously generated during cognitive processing or communication, possess representational qualities, and can influence human cognitive processing. Although researchers have varied emphases in their conceptual definitions of gesture, it is widely believed that gestures differ from direct actions and possess cognitive functions. Representative theoretical models of gesture' s cognitive functions include the Lexical Index Model, Information Packaging Hypothesis, Image Maintenance Theory, Semantic Specificity Hypothesis, and Embedded/Extended View. Based on differences in the primary independent variables in research on gesture' s cognitive functions, such research can be divided into three distinct paradigms: the gesture allowance/restriction paradigm, the gesture pattern modification paradigm, and the contextual variation paradigm. Future research directions worth pursuing include in-depth investigation of the neural mechanisms underlying gesture's cognitive functions, strengthening intervention studies on gesture' s cognitive functions, and establishing a more explanatorily powerful theoretical model of gesture' s cognitive functions—the 'spatialization' gesture hypothesis.

Full Text

A New Perspective on the Cognitive Function of Gestures: The "Spatializing" Gesture Hypothesis

Yu Wenhua; Lu Zhongyi
Hebei Normal University, Shijiazhuang 050024, China

Abstract: Human beings frequently produce gestures while speaking or thinking. Gestures arise automatically during cognitive processing or communication and serve a representational function, while also influencing human cognitive processing. Although researchers differ in their conceptual emphases regarding gestures, there is general consensus that gestures are distinct from direct actions and possess cognitive functions. Representative theoretical models of gesture' s cognitive function include the Lexical Index Model, Information Packaging Hypothesis, Image Maintenance Theory, Semantic Specificity Hypothesis, and Embedded/Extended Perspective. Based on the primary independent variables employed in gesture research, three distinct paradigms can be identified: allow-and-restrict gesture paradigms, gesture pattern modification paradigms, and context manipulation paradigms. Future research directions should focus not only on exploring the neural mechanisms underlying gesture' s cognitive functions and strengthening intervention studies, but also on proposing a more theoretically powerful model—the "Spatializing" Gesture Hypothesis.

Keywords: gestures; cognitive function; lexical index; information packaging; embodied cognition

---

2. Conceptual Definition of Gestures

Current debates persist regarding the conceptual definition of gestures (McNeill, 1985; Gentner, 1982; Hostetter & Alibali, 2008; Clark, 2013). Some research defines gestures functionally, viewing them as spontaneous hand movements that accompany speech (McNeill, 1992), structured representations (Gentner, 1982) that depict actions, movements, or shapes through hand form or motion (McNeill, 1985). These gestures are abstract representations detached from actual actions and objects, capable of more powerfully influencing thought and learning to achieve generalization and transformation of knowledge (Novack & Goldin-Meadow, 2016). Other studies adopt an embodied cognition perspective, considering gestures as an expression of embodiment (Hostetter & Alibali, 2008) and physical tools that extend the cognitive system, providing stable external physical and visual representations that furnish means for thinking (Clark, 2013). Some researchers have attempted to experimentally determine the conditions under which hand movements are perceived as gestures. Studies show that empty-hand movements are more readily perceived as gestures when target objects are present, the hand shape is grasp-like, and speech accompanies the motion (Novack, Wakefield & Goldin-Meadow, 2016).

Although gestures often accompany language, speech accompaniment is not essential for gesture recognition; environmental cues and top-down processing can achieve this recognition. Environmental cues include internal movement features (hand shape, path, rhythm, speed, etc.), external movement features (presence of target objects, prior knowledge and experience, etc.), and communicative cues (facial expressions, gaze, presence of listeners, etc.). Novack and Goldin-Meadow (2016) propose that gestures can be broadly categorized into representational and non-representational gestures. Representational gestures further divide into iconic gestures, metaphoric gestures, beat gestures, and deictic gestures (McNeill, 1992). Iconic gestures use hand and arm movements to present concrete objects or movement images. Metaphoric gestures employ space to convey abstract meanings. Beat gestures accompany speech rhythm to highlight important words. Deictic gestures point to specific locations of targets in the speaker' s environment. Lhommet and Marsella (2014) classify gestures into metaphoric, symbolic, iconic, and deictic types. Other scholars define descriptive and metaphoric gestures contextually, considering gestures in concrete contexts as descriptive and those accompanying abstract linguistic content as metaphoric (Straube, Green, Bromberger & Kircher, 2011).

Despite varying emphases in conceptual definitions, consensus exists on several points. First, gestures possess cognitive functions (Pouw, Nooijer, Gog, Zwaan & Paas, 2014). Gestures can reflect and influence individuals' cognitive processing (Mol et al., 2012; Goldin-Meadow & Wagner, 2005), significantly affecting cognitive activities such as language expression, creative thinking, and problem-solving. Gestures enhance listeners' comprehension and serve as an important source of sensorimotor information for both communication partners (Cook & Tanenhaus, 2009).

Second, gestures differ from direct actions. Although research confirms that action production potential or experience can increase gesture frequency (Pine, Gurney & Fletcher, 2010; Hostetter et al., 2010), and that hand movements are more readily perceived as gestures when target objects are present (Novack et al., 2016), gestures remain distinct from direct actions. Gestures do not directly manipulate objects; they are representational hand movements whose purpose is to indicate or represent other movements, objects, or even abstract concepts (Novack et al., 2016), possessing both concrete and abstract characteristics (Novack et al., 2014). Unlike direct actions bound by concrete goals, gestures enrich richer internal representations of targets, establishing stronger connections between bodily movement and thought as a form of schematized action. Gestures exert stronger effects on thinking than actual actions because they can more powerfully influence problem-solvers' representations of problems and more effectively facilitate encoding of spatial movement information (Kita, Alibali & Chu, 2017).

3. Theoretical Models of Gesture Cognitive Function

To explore the psychological mechanisms underlying gesture' s cognitive functions, researchers have proposed various theoretical models from different perspectives. Representative models include the Lexical Index Model, Information Packaging Hypothesis, Image Maintenance Theory, Semantic Specificity Hypothesis, and Embedded/Extended Perspective.

3.1 Lexical Index Model

The Lexical Index Model posits that gestures involve integration with the surface form of spoken language, where speakers' gestures can enhance lexical activation and facilitate semantic access. Gestures and lexical semantic components share common features that promote lexical selection or cross-modal representation (Krauss, Chen & Gottesmann, 2000). Research within this framework primarily focuses on speech-accompanying gestures, examining their impact on language fluency. The model predicts that restricting gesture use results in more speech pauses, slower speech rate, reduced fluency, and increased tip-of-the-tongue phenomena. Buccino et al. (2001) found that Broca's area is activated when making gestures. Evolutionary neuroscience evidence indicates evolutionary continuity between gesture and language, with monkey brain area F5 and human Broca' s area being homologous brain regions (Arbib, 2005). Area F5 is responsible for visuomotor coding of hand and mouth movements (Rizzolatti et al., 1988), with posterior F5 (F5p) playing a key role in producing and controlling hand movements (Raos, Umilta, Murata, Fogassi & Gallese, 2006; Umilta, Brochier, Spinks & Lemon, 2007), while Broca' s area is the primary brain region for human language processing. This suggests that Broca' s area (F5) may simultaneously control both gestures and word articulation.

3.2 Information Packaging Hypothesis

The Information Packaging Hypothesis proposes that gestures function in packaging spatial movement information into units expressible through language. Gestures help extract relevant features from scenes and package them into cognitive units (Gentner, 1982). Research demonstrates that more difficult information packaging activates more gestures (Kita, De & Mohr, 2007; Kita & Davies, 2009). Goldin-Meadow (2015) conducted cross-cultural gesture research with English and Turkish speakers, finding different gesture patterns under speech-accompanying conditions. English speakers showed a conflated pattern, combining manner and path into a single gesture, whereas Turkish speakers showed a separate pattern, using one gesture for path and another for manner, often producing only path gestures without manner gestures—consistent with linguistic characteristics. However, under non-speech conditions, both groups tended to simultaneously represent path and manner with a single gesture, demonstrating the existence of "natural semantic organization" when humans do not use language to convey motion events. Expression of motion events is driven by the need to transmit comprehensive information with limited effort, leading to information packaging tendencies in gestures. Mol and Kita (2012) instructed participants to represent manner and path with either the same gesture or different gestures, observing how this affected linguistic packaging of manner and path information. The study found that when participants used one gesture to simultaneously represent manner and path, they tended to express this information in a single linguistic sentence; when using separate gestures, they tended to use separate sentences. This demonstrates that how people gesture influences how they speak, supporting the Information Packaging Hypothesis. Moreover, manner word usage was highly correlated across both gesture conditions, indicating that different gesture patterns did not significantly affect lexical selection in speech, and that gestures do not influence language through lexical mechanisms.

3.3 Image Maintenance Theory

Image Maintenance Theory posits that gestures activate and maintain speakers' stored mental images that would otherwise rapidly fade. This theory reveals several functions of gestures. First, gestures can enrich listeners' mental images, particularly regarding spatial aspects. Vision is fundamentally hand-centered; seeing our hands and their positions is crucial when we move. Receiving erroneous visual information about hand position impairs our ability to make hand movements toward specific targets, placing the hand in a central role in representing movement direction (Loughlin, 2016). Restricting gestures reduces the imagery quality of language (Rimé, Bernard, Schiaratura, Hupet & Ghysselinckx, 1984). Second, gestures are excellent tools for capturing visual and spatial information, particularly suitable for describing spatial relationships such as hospital locations or neighborhood layouts. Gestures also accompany language content unrelated to spatial or visual imagery because people use hand movements and shapes to describe non-spatial information in visual and spatial forms, such as gestures used when explaining mathematical equations like 2+3=5, which convey mathematical relationships through spatial movement (Cook, Nusbaum & Goldin-Meadow, 2004). Therefore, while gestures are not necessary for lexical retrieval, they can facilitate it through imagery (Wesp, Hesse, Keutmann & Wheaton, 2001).

3.4 Semantic Specificity Hypothesis

People produce more gestures when thinking or speaking about movement-related content compared to other topics. The Semantic Specificity Hypothesis proposes that concrete, depictable, spatial, and manipulable semantic components elicit higher gesture rates. Pine et al. (2010) had participants perform object description tasks and found, through analysis of their speech and gestures, that descriptions of highly manipulable targets contained more iconic gestures regardless of listener presence. Semantic components containing high spatial movement content activated more gestures. Chu and Kita (2015) reached similar conclusions.

3.5 Embedded/Extended Perspective

Researchers holding the Embedded/Extended Perspective view gestures as physical tools that extend the cognitive system, providing external support for problem-solving and furnishing means for thought (Clark, 2013). They primarily focus on gestures produced during problem-solving, finding that when individuals face external constraints (high cognitive load, difficulty extracting information from the environment) or internal constraints (low working memory capacity), the cognitive system tends to select more accessible, externally supported problem-solving strategies, thereby producing gestures (Clark, 2008, 2013; Wheeler, 2013). Gestures can also facilitate categorization. Sixtus, Fischer and Lindemann had participants make number-indicating gestures or non-number-indicating gestures while classifying auditorily presented numbers. Only number-indicating gestures facilitated classification responses to the auditory numbers, while non-number-indicating gestures showed no such effect (Sixtus, Fischer & Lindemann, 2017).

4. Research Paradigms for Gesture Cognitive Function

Due to varying conceptual definitions, researchers employ different research paradigms. Based on the primary independent variables in gesture cognitive function research, existing paradigms can be divided into three categories.

4.1 Allow-Restrict Gesture Paradigm

This is the most commonly used paradigm for studying gesture' s cognitive function. Researchers randomly assign participants to groups that either encourage or allow gesture use versus groups that restrict gesture use, then have them complete cognitive tasks to measure gesture' s impact on task performance. Beaudoin-Ryan and Goldin-Meadow (2014) randomly assigned fifth-grade students to encourage-gesture, discourage-gesture, or control conditions for Kohlberg' s moral dilemma tasks, finding significantly enhanced perspective-taking ability in the encourage-gesture group. In a route memory task, participants who could silently simulate the route with gestures showed significantly better recall than those who merely observed while holding balls in both hands to restrict gestures (So, Ching, Lim, Cheng & Lp, 2014). Cook et al. (2004) had participants memorize linguistic and visuospatial materials under gesture-allowed versus gesture-restricted conditions. Participants remembered more items in both memory domains when gestures were allowed, demonstrating that gestures significantly facilitate not only verbal working memory but also visuospatial working memory. Additionally, whether gestures are allowed can influence the type of information and strategy selection relied upon in problem-solving (Alibali et al., 2011).

4.2 Gesture Pattern Modification Paradigm

This paradigm examines how different gesture patterns affect cognitive activities to study gesture' s cognitive function. Mumford and Kita (2013) used different gestures to teach three-year-old children novel words representing state changes or manners. Children who learned through manner gestures showed stronger manner preferences, demonstrating that specific gesture forms guide children' s interpretation of novel words. Mol et al. (2012) studied how different gesture patterns affect syntactic structure in speech. Other researchers have examined how finger grasping or opening actions affect numerical extraction (Grade, Badets & Pesenti, 2016). Leung et al. (2012) had participants generate novel uses for a campus building while performing a speech gesture (extending one hand and arm palm-up). The experiment consisted of two trials with identical questions (unknown to participants). Experimental group participants switched hand positions between trials (right hand up/left hand behind back vs. left hand up/right hand behind back), while control participants maintained the same posture across both trials. Switching hand positions enhanced divergent thinking fluency, flexibility, and originality. Novack et al. (2014) investigated the effects of concrete versus abstract gestures on learning, finding that abstract gestures in teaching more effectively promoted generalized learning of knowledge compared to concrete gestures.

4.3 Context Change Paradigm

This paradigm examines how background factors such as task difficulty, communicative context, and cultural differences affect gesture selection. When participants explained complex scientific system diagrams to experts versus novices, they produced more descriptive gestures for novices and more deictic gestures for experts, packaging more information for novices through multiple channels (Kang, Tversky & Black, 2015). Other research has focused on the relationship between task difficulty and gesture frequency (Chu & Kita, 2008, 2011; Kita et al., 2007; Kita et al., 2009; Hostetter et al., 2011), finding that more difficult tasks elicit more gestures. Chu et al. (2015) studied how target manipulability affected gestures in mental rotation and motion event description tasks, finding significantly higher gesture frequency in the high-manipulability condition, consistent with Pine et al. (2010).

5.1 Summary of Existing Theories and Research

Researchers have attempted to establish theoretical models explaining the generation and mechanisms of gesture's cognitive functions from various perspectives. However, these models typically explain only certain gesture types or partial mechanisms, failing to provide comprehensive explanations across broader dimensions.

5.1.1 Limited Explanatory Power for Gesture Cognitive Function

Existing theories and research have focused more on gesture's communicative functions, such as facilitating communication, influencing language, and promoting sensorimotor information exchange. The Lexical Index Model addresses gesture's role in semantic access, partially explaining the relationship between gesture and language fluency and demonstrating some correlation between gesture and language, but cannot effectively explain mechanisms at more complex levels such as linguistic structure and grammatical features, nor mechanisms in non-verbal tasks. The Information Packaging Hypothesis suggests gestures package spatial movement information into units suitable for linguistic expression, reducing cognitive load. Research within this framework primarily focuses on relationships between gesture patterns and linguistic expression. Although examining larger units than lexical items, it fails to clearly explain key issues such as how information is packaged and how non-spatial movement information is bundled. Image Maintenance Theory primarily addresses gesture's importance in enhancing the imagery quality and maintenance duration of information processed in working memory, but lacks strong explanations for gesture functions in other cognitive processes and activities. The Semantic Specificity Hypothesis notes the correlation between spatial, manipulable semantic components and higher gesture rates, proposing that specific semantic components unidirectionally activate gestures. This theory cannot effectively explain gesture' s effects on language expression, problem-solving, creative thinking, and learning. The Embedded/Extended Perspective views gestures as external physical extensions of the cognitive system with implicit connections to spatial locations. While reasonable, research within this framework focuses more on gesture' s embodied nature than on its cognitive functions and implementation mechanisms.

Kita et al. (2017) attempted to establish a new theoretical framework in their review—the Gestures-for-Conceptualization Hypothesis. This hypothesis explains gesture' s self-directed function, proposing that gestures influence cognitive processing through four pathways: activating, manipulating, packaging, and exploring spatial movement information. However, this theory only addresses gesture' s representational nature and self-directed function, and lacks supporting experimental evidence.

Regarding research paradigms, the allow-restrict paradigm primarily examines effects of gesture presence versus absence and spontaneous gestures on cognitive activities. The gesture pattern modification paradigm investigates how different gesture patterns (hand shape, orientation, path) affect cognitive activities. The context change paradigm treats gesture as a dependent variable, examining how task difficulty, communicative context, and cultural differences affect gesture selection. Moreover, most existing paradigms are behavioral, revealing gesture's effects on language and other cognitive activities from different perspectives but unable to deeply analyze internal cognitive neural structures and mechanisms.

5.1.2 Need for Deeper Research on Embodied Nature of Metaphoric Gestures

Over decades, psychologists have conducted extensive research on the embodied nature of language comprehension. This includes studies on perceptual involvement during language comprehension (Kaup, Yaxley, Madden, Zwaan, & Lüdtke, 2007; Connell, 2007; Zhang & Lu) and substantial evidence showing activation and simulation of actions, particularly hand actions, during language comprehension. For example, with the verb "push-pull," Glenberg and Kaschak (2002) found participants responded faster to targets moving away from the body when reading "close the drawer" and faster to approaching targets for "open the drawer." With rotation verbs, participants rotated clockwise faster when hearing "tighten the screw" (Zwaan & Taylor, 2006). Tucker and Ellis (2001) found that when judging whether objects like "key" and "potato" were natural, small objects activated fine "pinching" actions while large objects activated whole-hand "grasping" actions. These studies represent advances over traditional language comprehension research. As Glenberg and Gallese (2012) noted, linguistic symbols become more meaningful through perceptual, motor, and emotional systems.

Two points emerge from this research: First, embodied language comprehension studies have primarily focused on concrete concepts and their literal meanings. Second, hand actions during language comprehension involve simple embodied responses or movements. In contrast, research on embodied cognition of abstract concepts, particularly metaphorical concepts, remains insufficiently systematic and in-depth. Future research extending from literal meanings of concrete concepts to figurative meanings of abstract concepts, and from hand actions containing concrete movements to representational linguistic gestures, will undoubtedly advance embodied language comprehension research.

Metaphoric gestures can present abstract meanings using space. For example, when expressing "grasping an idea," the accompanying gesture typically involves a cupped hand shape as if holding something, with the gesture functioning like a container—this is a conduit metaphor gesture. Conduit metaphor gestures can also illustrate relationships between concepts by placing represented concepts in space (Hasegawa, Shirakawa, Shioiri, Hanawa & Ohara, 2015). Compared to other types, metaphoric gestures hold greater significance for studying gesture cognitive functions, particularly higher-level cognitive functions, due to their close connection with abstract concepts and semantic representation. Researchers have confirmed implicit connections between gestures and spatial locations from an embodied cognition perspective, finding implicit conceptual associations between the index finger and "top" and the thumb and "bottom" (Romano, Marini & Maravita, 2017). Additionally, finger grasping actions can affect the magnitude of numbers extracted in numerical tasks, with grasping extracting smaller numbers and finger opening extracting larger numbers (Grade et al., 2016). However, current research on the embodied nature of metaphoric gestures remains insufficiently systematic and in-depth, lacking thorough investigation into theoretical models, neural mechanisms, and interventions.

5.2.1 Establishing a More Explanatory Theoretical Model of Gesture Cognitive Function

Through comprehensive analysis of existing theoretical and experimental research on gesture' s cognitive functions, we find that gestures can activate and maintain spatial imagery, facilitate access to and communication of spatial movement information, achieve "spatialization" of non-spatial information, and thereby ease selection of perceptual strategies. Therefore, from an embodied cognition perspective and focusing on metaphoric gestures' cognitive functions, we propose the "Spatializing" Gesture Hypothesis.

The core proposition of the Spatializing Gesture Hypothesis is that abstract concept understanding is grounded in the sensorimotor experience of hands in space, using hand shape, path, position, and other features to highlight, emphasize, and organize spatial movement information. This effectively transforms non-spatial movement information into spatial movement information, enhances visual or kinesthetic representation of information, reduces cognitive load, and facilitates problem-solving. For example, using both hands to represent two different possibilities in problem-solving can promote multiple representations of the problem, and switching from palm-up to palm-down can facilitate thinking about opposing viewpoints. Thus, from the embodied cognition perspective, gestures reflect encoding of bodily sensory and motor information. According to our Spatializing Gesture Hypothesis, when humans think or talk about perception, imagery, and situations, they map primary metaphorical abstract features —such as time, causality, morality, and emotion—onto concrete physical features to form gestures according to different communicative intentions.

This mapping mechanism operates in two primary ways: listing and contrasting (Lhommet et al., 2014). Listing involves sequentially describing objects using gestures, such as using the right index finger to sequentially count the left hand' s fingers to represent four aspects of something. Contrasting can be viewed as an axis metaphor, displaying differences between two or more objects by converting each into a scalar value at different positions along an axis. Axes can be horizontal, vertical, or front-back. Horizontal and front-back axes primarily represent time metaphors. When using the horizontal axis, the subject is typically not involved, with left representing past, center representing present, and right representing future for right-handed individuals. When using the front-back axis, the subject is involved, with front representing future and back representing past. The vertical axis uses gesture' s vertical spatial position to map abstract concepts: more is up, less is down; moral is up, immoral is down; positive emotions are up, negative emotions are down, etc. Through mapping from abstract to concrete features, gestures facilitate activation (So et al., 2014; Hostetter et al., 2008, 2010), maintenance (Romano et al., 2017), and communication (Buccino et al., 2001; Cook et al., 2009; Pine et al., 2010; Clark, 2013; Mol et al., 2012; Chu et al., 2015) of spatial movement information, easing its representation. Simultaneously, gestures facilitate "spatialization" of non-spatial movement information, enabling more effective access (Alibali et al., 2011; Goodrich & Hudson, 2009; Goldin-Meadow, Nusbaum, Kelly & Wagner, 2001; Grade et al., 2016; Sixtus et al., 2017), packaging (Kita et al., 2007; Kita et al., 2009; Gentner, 1982), and organization (Beaudoin-Ryan et al., 2014; Chu et al., 2008, 2011; Kirk et al., 2017; Leung et al., 2012) of non-spatial information. In summary, gestures reduce perceptual load and facilitate information processing.

This hypothesis possesses strong explanatory power, but it is built upon existing research reviews and theoretical proposals. The next step requires experimental research to obtain more robust empirical support.

5.2.2 Deepening Research on Neural Mechanisms of Gesture Cognitive Function

Most existing research on gesture' s cognitive functions employs behavioral experiments, with only occasional attempts to use functional magnetic resonance imaging (fMRI) to examine underlying physiological mechanisms (Goldin-Meadow et al., 2001; Buccino et al., 2001), resulting in limited studies. Moreover, most current research uses offline analysis, leading to low temporal resolution in analyzing gesture' s cognitive functions and compromising result validity and reliability. Future research should employ functional magnetic resonance imaging, event-related potentials, and other techniques to real-time investigate the neural mechanisms of gesture' s cognitive functions.

5.2.3 Strengthening Intervention Research on Gesture Cognitive Function

Gestures accompany human speech and thought processes, enabling generalization and transformation of knowledge. Gestures possess powerful cognitive functions. Compared to actual actions, gestures more powerfully promote human cognition and thinking (Novack et al., 2016; Kita et al., 2017). Numerous studies confirm that gestures influence human cognitive processing, even significantly affecting problem-solving (Chu et al., 2015; Hostetter et al., 2010; Alibali et al., 2011) and creative thinking (Beaudoin-Ryan et al., 2014; Leung et al., 2012; Kirk et al., 2017). However, current research remains largely at the basic stage, with very few intervention studies utilizing gestures to promote human learning (Novack et al., 2014), thinking, and creativity, which greatly limits effective application of gesture' s cognitive functions. Future research should emphasize and strengthen intervention studies on gesture' s cognitive function, designing gesture-related intervention programs based on existing findings to further enhance the applied value of gesture patterns.

References

Zhang, P., & Lu, Z. Y. (2013). Mental representation of color information in sentence comprehension. *Acta Psychologica Sinica*, 45(4), 406−415.

Alibali, M. W., Spencer, R. C., Knox, L., & Kita, S. (2011). Spontaneous gestures influence strategy choices in problem solving. *Psychological Science*, 22(9), 1138−1144.

Arbib, M. A. (2005). From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28(2), 105−124.

Beaudoin-Ryan, L., & Goldin-Meadow, S. (2014). Teaching moral reasoning through gesture. *Developmental Science*, 17(6), 984-990.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., & Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, 13(2), 400−404.

Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General*, 137, 706-723.

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology General*, 140(1), 102−116.

Chu, M., & Kita, S. (2015). Co-thought and co-speech gestures are generated by the same action generation process. *Journal of Experimental Psychology Learning Memory & Cognition*, 42(2), 257−270.

Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York, NY: Oxford University Press.

Clark, A. (2013). Gesture as thought. In Z. Radman, *The Hand, an Organ of the Mind: What the Manual Tells the Mental* (pp.255-268). Cambridge, Britain: MIT Press.

Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, 102(3), 476−485.

Cook, S. W., Nusbaum, H., & Goldin-meadow, S. (2004). Probing the mental representation of gesture: is handwaving spatial?. *Journal of Memory & Language*, 50(4), 395−407.

Cook, S. W., & Tanenhaus, M. K. (2009). Embodied communication: speakers' gestures affect listeners' actions. *Cognition*, 113(1), 98−104.

Gentner, D. (1982). Why nouns are learned before verbs: linguistic relativity versus natural partitioning. Technical report no. 257. *Child Language*, 2, 86.

Glenberg, A. M., & Gallese, V. (2012). Action-based Language: A theory of language acquisition, comprehension, and production. *Cortex*, 48, 905-922.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9(3), 558−565.

Goldin-meadow, S. (2000). Beyond words: the importance of gesture to researchers and learners. *Child Development*, 71(1), 231.

Goldin-meadow, S. (2015). Does language shape silent gesture?. *Cognition*, 148, 10−18.

Goldin-Meadow, S., & Wagner, S. M. (2005). How our hands help us learn. *Trends in Cognitive Sciences*, 9(5), 234−241.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: gesturing lightens the load. *Psychological Science*, 12(6), 516−522.

Goodrich, W., & Hudson Kam, C. L. (2009). Co-speech gesture as input in verb learning. *Developmental Science*, 12(1), 81—87.

Grade, S., Badets, A., & Pesenti, M. (2016). Influence of finger and mouth action observation on random number generation: an instance of embodied cognition for abstract concepts. *Psychological Research*, 1—11.

Hasegawa, D., Shirakawa, S., Shioiri, N., Hanawa, T., & Ohara, K. (2015). The effect of metaphoric gestures on schematic understanding of instruction performed by a pedagogical conversational agent. *Lecture Notes in Computer Science*, 361—371.

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: gestures as simulated action. *Psychonomic Bulletin & Review*, 15(3), 495—514.

Hostetter, A. B., & Alibali, M. W. (2010). Language, gesture, action! a test of the gesture as simulated action framework. *Journal of Memory & Language*, 63(2), 245—257.

Kang, S., Tversky, B., & Black, J. B. (2015). Coordinating gesture, word, and diagram: Explanations for experts and novices. *Spatial Cognition & Computation*, 15(1), 1—26.

Kaup, B., Yaxley, R. H., Madden, C. J., Zwaan, R. A., & Lüdtke, J. (2007). Experiential simulations of negated text information. *Quarterly Journal of Experimental Psychology*, 60(7), 976-990.

Kirk, E., & Lewis, C. (2017). Gesture facilitates children' s creative thinking. *Psychological Science*, 28(2), 225—232.

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, 124(3), 245—266.

Kita, S., De, C. O., & Mohr, C. (2007). Metaphor explanation attenuates the right-hand preference for depictive co-speech gestures that imitate actions. *Brain & Language*, 101(3), 185—197.

Kita, S., & Davies, T. S. (2009). Competing conceptual representations trigger co-speech representational gestures. *Language & Cognitive Processes*, 24(5), 761—775.

Krauss, R. M., Chen, Y., & Gottesmann, R. F. (2000). Lexical gestures and lexical access: a process model. In D. McNeill, *Language and Gesture* (pp.261—283). Cambridge, Britain: Cambridge University Press.

Leung, A. K., Kim, S., Polman, E., Ong, L. S., Qiu, L., & Goncalo, J. A., et al. (2012). Embodied metaphors and creative "acts" . *Psychological Science*, 23(5), 502—509.

Lhommet, M., & Marsella, S. (2014). Metaphoric Gestures: Towards Grounded Mental Spaces. 14th International Conference on Intelligent Virtual Agents

Conference. Berlin, Germany: Springer International Publishing.

Loughlin, V. (2016). Zdravko radman, the hand: an organ of the mind, what the manual tells the mental. *Phenomenology & the Cognitive Sciences*, 15(2), 291−296.

Mcgregor, K. K., Rohlfing, K. J., Bean, A., & Marschner, E. (2009). Gesture as a support for word learning: the case of under. *Journal of Child Language*, 36(4), 807−828.

Mcneill, D. (1985). So you think gestures are nonverbal?. *Psychological Review*, 92(3), 350−371.

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

Mol, L., & Kita, S. (2012). Gesture structure affects syntactic structure in speech. *Conference of the Cognitive Science Society. Cognitive Science Society*, 34(34), 761−766.

Mumford, K. H., & Kita, S. (2013). Children use gesture to interpret novel verb meanings. *Child Development*, 85(3), 1181−1189.

Novack, M. A., & Goldin-Meadow, S. (2016). Gesture as representational action: A paper about function. *Psychonomic Bulletin & Review*, 24(3), 1−14.

Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: using the hands to learn math. *Psychological Science*, 25(4), 903−910.

Novack, M. A., Wakefield, E. M., & Goldin-Meadow, S. (2016). What makes a movement a gesture?. *Cognition*, 146, 339−348.

Pine, K. J., Gurney, D. J., & Fletcher, B. (2010). The semantic specificity hypothesis: when gestures do not depend upon the presence of a listener. *Journal of Nonverbal Behavior*, 34(3), 169−178.

Pouw, W. T. J. L., Nooijer, J. A. D., Gog, T. V., Zwaan, R. A., & Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Frontiers in Psychology*, 5(5), 1−14.

Raos V, Umilta M-A, Murata A, Fogassi L, Gallese V. (2006). Functional properties of grasping-related neurons in the ventral premotor area F5 of the macaque monkey. *Journal of Neurophysiology*, 95(5), 709-729.

Rimé, Bernard, Schiaratura, L., Hupet, M. & Ghysselinckx, A. (1984). Effects of relative immobilization on the speakers nonverbal behavior and on the dialog imagery level. *Motivation & Emotion*, 8(4), 311−325.

Rizzolatti G, Camarda R, Fogassi M, Gentilucci M, Luppino G, Matelli M. (1988). Functional organization of inferior area 6 in the macaque monkey: II. Area F5 and the control of distal movements. *Experimental Brain Research*, 71(3), 491-507.

Romano, D., Marini, F., & Maravita, A. (2017). Standard body-space relationships: fingers hold spatial information. *Cognition*, 165, 105—112.

Sixtus, E., Fischer, M. H., & Lindemann, O. (2017). Finger posing primes number comprehension. *Cognitive Processing*, 1—12.

So, W. C., Ching, T. H., Lim, P. E., Cheng, X., & Lp, K. Y. (2014). Producing gestures facilitates route learning. *Plos One*, 9(11), 1—21.

Straube, B., Green, A., Bromberger, B., & Kircher, T. (2011). The differentiation of iconic and metaphoric gestures: common and unique integration processes. *Human Brain Mapping*, 32(4), 520—533.

Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition*, 8, 769—800.

Umilta MA, Brochier T, Spinks RL, Lemon RN. (2007). Simultaneous recording of macaque premotor and primary motor cortex neuronal populations reveals different functional contributions to visual motor grasp. *Journal of Neurophysiology*, 98(1), 488-501.

Wesp, R., Hesse, J., Keutmann, D., & Wheaton, K. (2001). Gestures maintain spatial imagery. *American Journal of Psychology*, 114(4), 591—600.

Wheeler, M. (2013). Is cognition embedded or extended? The case of gestures. In Z. Radman, *The Hand, an Organ of the Mind: What the Manual Tells the Mental* (pp.269-301). Cambridge, Britain: MIT Press.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, 135, 1-11.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.