158. Embodied meaning, inside and out: The coupling of gesture and mental simulation

Abstract

During situated interaction, meaning construction is embodied inside and out. Meaning is embodied internally when we create embodied simulations, co-opting brain areas specialized for perception or action to create dynamic mental representations, rich in sensorimotor detail. Thinking of petting a kitten, for instance, might include visual simulation of its appearance and motor simulation of the act of petting – all in brain areas typically used to see and touch kittens. At the same time, meaning is embodied externally in representational gestures, actions of the hands and body that represent objects, actions, and ideas. In this chapter, we argue that these internal and external embodiments are tightly coupled, with bidirectional causal influences between gesture and simulation, both within the speaker and between speaker and listener. By embodying meaning, inside and out, language users take advantage of the complementary semiotic affordances of thought and action, brain and body.

1. Inside and out: Mental simulation and gesture

The creation of meaning is intimately tied to the body. Or rather, bodies. Meaning-making is seldom solitary, and the prototypical linguistic encounter involves multiple interacting agents working together to negotiate shared meaning. These multiple interlocutors use their bodies as fully fleshed out semiotic resources during situated talk (Goodwin 2000; Kendon 2004; McNeill 2005). That is, they *gesture*, moving their hands and bodies in meaningful ways. Meaning is also embodied less visibly. The neural activity supporting language comprehension and production relies on brain areas that are re-purposed from perception and action – that is, "embodied" brain areas – and these areas coordinate during comprehension to create "embodied simulations" of linguistic content (Barsalou 2008; Glenberg 2010). In this chapter, we argue that simulation and gesture are forms of embodiment that are tightly and multiply coupled during situated meaning-making.

To start, the body makes an invisible contribution to meaning in virtue of interlocutors' embodied simulations (Barsalou 2008). When comprehenders hear language describing bodily actions, their brains' motor systems become engaged in a content-specific way. That is, hearing a sentence about kicking engages motor circuitry responsible for controlling leg actions, just as hearing a sentence about chewing lights up neural areas for actual mouth-moving (Pulvermüller and Fadiga 2010). This content-specific use of

Müller, Cienki, Fricke, Ladewig, McNeill, Teßendorf (eds.) 2014, Body – Language – Communication (HSK 38.2), de Gruyter, 2000–2007

brain circuits specialized for action is paralleled in perception: processing language about visible objects and events activates visual regions of the brain (Stanfield and Zwaan 2001), while language about audible things prompts a detailed auditory simulation of their sound (Winter and Bergen 2012). There is now convergent evidence for embodied simulation during language comprehension from brain imaging, behavioral experimentation, and neuropsychological lesion studies (Bergen 2012). This internal simulation of what it might be like to perform described actions or perceive described objects and events has been most thoroughly studied during language comprehension, but there's initial evidence that it plays a role in language production as well. One leading idea (Sato 2010) is that speakers perform embodied simulations at the phase of message generation (cf. Levelt 1993) during language production. If a message includes percepts or actions, embodied simulation might play a role in its construction.

A second, more visible way in which the body is recruited to create meaning is through the production of gestures, spontaneous movements – typically of the hands – that take advantage of the motor-visual channel to complement speech. Representational gestures, in particular, use a combination of hand morphology, motion trajectory, and spatial location to represent concrete or abstract referents through literal or metaphoric resemblance. While speech and gesture are tightly coupled in both timing and meaning, they often express information that is complementary rather than identical (McNeill 2005). You might ask, "Have you grown?" while using an extended index finger to trace an upward trajectory – communicating that the question is about height, not heft. Or you could say, "My mood has really changed," while pointing upward, using space metaphorically to communicate that your mood has changed for the better.

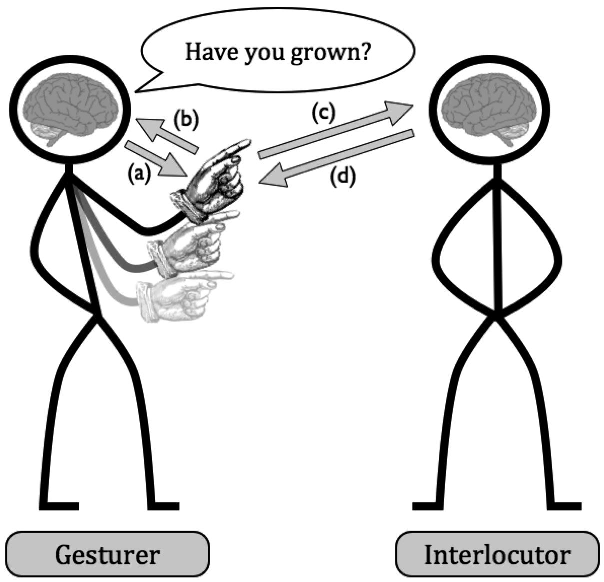

Fig. 158.1: Relations between gesture and simulation: speaker simulation drives speaker gesture (a); speaker gesture influences speaker simulation (b) and listener simulation (c); and listener simulation shapes their interpretation of speaker gesture (d).

Situated meaning-making is embodied inside and out: in the internal embodied simulations that accompany message formulation and comprehension, and in the external meaningful gestures that accompany speech. Crucially, these embodied processes work in concert (Fig. 158.1): a speaker's mental simulation drives her gestures (§3); her gestures influence her own (§4) and listeners' simulations (§5); and a listener's ongoing simulation may even shape her interpretation of speakers' gestures (§6).

2. Speaker simulation shapes speaker gesture

Embodied simulation and gesture are both particularly well suited to representing action. Motor simulation re-engages neural circuitry that is also responsible for action performance; gesture, as meaningful manual *action*, can (re-)enact actions with high fidelity. Pursuing this similarity, Hostetter and Alibali (2008) argued that external gestures are actually the product of internal action simulation (see Fig. 158.1a). On their account, whenever motor or visual simulation of action exceeds a threshold of activation, the simulated action is expressed as real, gestural action. A number of other authors have suggested that gestures are the product of imagistic or motor representations generated during language production (e.g. Kita and Özyürerk 2003; McNeill 2005), although none of these have explicitly connected those imagistic processes to embodied neural simulation. This threshold may be sensitive to a number of factors, including neural connectivity, cognitive load, and socio-communicative norms. They suggest further that speech-related activation of the motor system makes simulated action most likely to "spill out" as gesture during language production, although gestures *are* sometimes produced in the absence of speech (e.g. Chu and Kita 2011).

In line with this proposal, several recent studies have found that increased action simulation predicts increased representational gestures. In Hostetter and Alibali (2010), for instance, participants either created or viewed geometric patterns, and then described them from memory. Participants produced more representational gestures when describing patterns that they had created themselves rather than merely viewed. Similarly, Sassenberg and Van Der Meer (2010) reported that, during route descriptions, representational gestures were more common during increased mental simulation, independent of task difficulty.

Gesture also reflects specific details of ongoing simulation. When people describe how they solved the Tower of Hanoi puzzle, which involves rearranging different sized disks, the trajectories of their gestures reflect the trajectories of the manual actions they actually used to solve the puzzle (Wagner Cook and Tanenhaus 2009), suggesting that gesture form is shaped by the particulars of motor simulation. Embodied simulation, moreover, can vary in focus, perspective, length, and degree of detail, and grammatical choices during language production may reflect these features of simulation. For example, producing progressive language may reflect increased focus in simulation on the ongoing process of the described state or event, as contrasted, say, with the resulting endstate (Bergen and Wheeler 2010; Madden and Zwaan 2003). And indeed when speakers produce progressive rather than perfect sentences, their accompanying gestures are longer-lasting and more complex (Parrill, Bergen, and Lichtenstein 2013). Similarly, during event descriptions, the viewpoint adopted in gesture − either internal or external to the event − is shaped by the body's involvement in the described event (Parrill 2010) and by the kind of motion information encoded in speech (Parrill 2011), both of which may reflect aspects of simulated viewpoint (Brunyé et al. 2009). To the extent that aspect and other grammatical features accurately diagnose properties of embodied simulation, it appears that speakers express the product of simulation in not only the linguistic channel, but also the gestural channel.

3. Speaker gesture shapes speaker simulation

A speaker's gestures can also prompt and shape their own embodied simulation (Fig. 158.1b). Representational gestures have been proposed to help maintain mental imagery during language production (de Ruiter 2000), or to "help speakers organize rich spatio-motoric information into packages suitable for speaking" (Kita 2000: 163). More generally, producing representational gestures may encourage an embodied simulation of the objects and events represented in gesture. For instance, Alibali and colleagues (2011) had participants solve variants of the following gear-rotation puzzle: If there are five gears in a row, each interlocked with its neighbors, and the first gear turns counterclockwise, then in what direction will the last gear turn? This puzzle can be solved using either a "sensorimotor" strategy that relies on a mentally simulating the gears' rotation or a more formal "parity" strategy that takes advantage of the alternating pattern of gear rotation (i.e. odd gears will turn counterclockwise, even gears clockwise). Gesturing made participants more likely to adopt the sensorimotor strategy, compared to trials where they spontaneously chose not to gesture and to trials where gesture was inhibited. Gesturing also improves performance on classic mental rotation tasks (Chu and Kita 2011), but only if the gesture is actually produced and not merely seen (Goldin-Meadow et al. 2012). Producing gestures, therefore, may encourage embodied simulation during problem solving.

Gesture production can also evoke detailed sensorimotor information. In one study, Beilock and Goldin-Meadow (2010) had participants solve the Tower of Hanoi puzzle (pre-test) and then solve it again after a brief pause (post-test). For half the participants, the relative weights of the disks were surreptitiously switched before the post-test, which impaired performance – but only if, between pre-test and post-test, participants explained how they had solved the puzzle in the pre-test. The fact that participants were unaffected by the switched weights if they hadn't explained their solution – and thus hadn't gestured about the task – suggests that gesture itself was responsible for shaping speakers' embodied representation of the task. Indeed, the impact of switching weights was mediated by the kinds of gestures produced during the explanation: the more the gestures represented the relative weights of the disks, the more post-test performance was affected by the switched weights. In other words, gesturing about the disks' task-irrelevant weights added precise weight information to participants' representations of the puzzle. These multimodal influences of gesture on thought suggest that gestures are not merely visible (Hofstetter and Alibali 2008) but also a distinctly felt form of embodiment, shaping the precise sensorimotor information included in accompanying simulation (cf. Goldin-Meadow et al. 2012).

4. Speaker gesture shapes listener simulation

One of the classic findings in the study of co-speech gesture is that speakers' gestures contribute to listeners' comprehension, affecting representations in long-term memory (e.g., Rogers 1978). One mechanism for this influence may be the modulation of listeners' motor or visual simulation (Fig. 158.1c), and several recent studies support this proposal.

Viewing gestures that represent actions may prompt a listener to simulate the performance of that action. Gestures that encode fine details of the trajectory of a manual action can lead the listener to include those details in their subsequent reproduction of the action, as if viewing the gesture prompted a detailed motor simulation of the action (Cook and Tanenhaus 2009). Moreover, a speaker's gestures may guide the listener's creation of a detailed visual simulation. In one experiment, Wu and Coulson (2007) showed participants videos of a man describing everyday objects, followed by a picture of the mentioned object. Crucially, the shape of the pictured object could be compatible with both the man's speech and gesture, or with his speech alone. For example, when the man said, "It's actually a double door," this could refer to Dutch-style double doors in which two half-doors are stacked one on top of the other, or French-style double doors where two full doors are placed side by side. When his speech was accompanied by a bimanual gesture that evoked Dutch-style double doors − his hands stacked on top of each other − then a subsequent picture of Dutch-style double doors would be related to both his speech and his gesture, while a picture of French-style double doors would be related to his speech but not his gesture. They found that pictures were easier to process when they matched both speech *and* gesture, as if the fine details of the man's gesture shaped the listener's visual simulation of the speech content. Gestures produced by a speaker during discourse, therefore, can rapidly shape the fine details of a listener's motor and visual simulation of discourse content.

5. Listener simulation shapes perception of speaker gesture

Any particular representational gesture can, in principle, represent an infinite range of hypothetical referents. Consider a gesture tracing an upward trajectory. Does it represent the flightpath of a bumblebee? Or a disco-dancing move? The precise meaning of a gesture is often disambiguated by discursive context or concurrent speech (e.g. "It flew to the flower" vs. "He danced like Travolta."). Another possible source of disambiguation, currently underexplored, is the comprehender's embodied simulation at the moment they perceive a gesture. A comprehender's concurrent or preceding simulation may shape their interpretation of a speaker's gestures, perhaps by selectively focusing attention on specific facets of a gesture's motor or visuospatial properties (Fig. 158.1d).

Consider the different ways in which a gesture can stand in for a referent: by *depicting*, tracing an object's shape in the air; by *enacting*, recreating some part of an action; or by *modeling*, using the hand to stand in for another object (Kendon 2004; cf. Müller 2009, Streeck 2008). A single gesture can be ambiguous among these. Consider a hypothetical gesture produced while saying, "I wiped the water off the table," in which a face-down open-palm handshape sweeps across a horizontal plane. This gesture could be *depicting* the tabletop's shape, height, or size; *enacting* the manual action of wiping; or using the hand to *model* the cloth itself. We hypothesize that the listener's interpretation of such a gesture, and thus the gesture's contribution to ongoing comprehension, may be shaped by their embodied simulation at the time of gesture perception. If the preceding linguistic context has prompted a motor simulation of the actions involved in cleaning ("My arm was so tired from wiping down the tabletop."), then the ongoing motor simulation at the moment of gesture perception might lead the listener to interpret the gesture as an *enactment*. In contrast, if the conversation up until then has focused on the table's appearance ("It was a gorgeous antique table with the flattest top you've ever seen, but there was water on it."), then the ongoing visual simulation of the table's appearance could increase attention to the way the gesture represents the visuospatial properties of the table or spill (e.g. its height or shape), prompting a *depictive* interpretation. Similarly, the viewpoint of a listener's ongoing simulation, either internal or external to an event, may influence their interpretation of gestures that are ambiguous with respect to viewpoint. The perceived meaning of a speaker's gesture, therefore, could be shaped by properties of the interlocutor's ongoing embodied simulation, although at present we know of no empirical support for this proposal.

6. Conclusion: Simulation, gesture, and the embodiment of meaning

Situated meaning-making involves the body in at least two ways: as gesture and as embodied simulation. As we have seen, these two sources of embodiment, external and internal, are tightly coupled both within and between interlocutors (Fig. 158.1).

This chapter has focused on concrete meanings: physical actions and objects, spatial arrays, perceivable events. But the human semantic potential far outstrips the concrete, allowing us to communicate about things unseen, displaced in time and space, and things *unseeable* like time, love, and mathematics. Even these abstract domains, however, may become meaningful through our bodies. Concepts as abstract as time and arithmetic may be mapped metaphorically to more concrete domains, and thus rely on sensorimotor simulations of those concrete source domains (Gibbs 2006; Lakoff and Johnson 1980). Gestures about abstract concepts that are understood metaphorically, moreover, often reflect the concepts' spatial and embodied sources (Cienki and Müller 2008; Marghetis and Núñez 2013). For instance, even though numbers do not literally vary in size, talk of "bigger" or "smaller" numbers may involve both spatial simulations of size and gestures that reflect spatial metaphors for number (Núñez and Marghetis to appear).

The fact that one source of embodiment is tucked inside the skull and the other is cast out into the world has implications for the representational work that they can do. First, gestures are visible to interlocutors, both intended and unintended, and are thus distinctly *public*; simulation, on the other hand, is shielded from others by the braincase, and is thus a *private* form of embodiment. Second, simulation is multimodal and thus generates a richly embodied representation, including taste and smell (e.g. Louwerse and Connell 2012); gesture, tied to the manual modality, forces a schematization of the represented content, tied more to action than to any other modality (Goldin-Meadow et al. 2012). Third, simulation isn't limited by the physiology of the hand (or body), and so it can represent the impossible, the difficult, the novel − but it is softly constrained by physiology and experience, so that it is more difficult to simulate unnatural or difficult actions (e.g. Flusberg and Boroditsky 2011). The body proper, by contrast, is perfectly suited for representing *itself*, without recourse to internal representations of its physiology; manual gestures afford rich representations of manual actions. In sum, gesture and simulation differ in terms of their publicity, their multimodal richness, and their representational affordances.

"It is my body which gives meaning," wrote Merleau-Ponty, "not only to the natural object, but also to cultural objects like words" (2002: 273). This meaning-giving body, we have argued, makes its contribution in at least two ways – as internal simulation, and as external gestures – each shaping the other during situated interaction.

7. References

Alibali, Martha W., Robert C. Spencer, Lucy Knox and Sotaro Kita 2011. Spontaneous gestures influence strategy choices in problem solving. *Psychological Science* 22(9): 1138–1144.

Barsalou, Lawrence W. 2008. Grounded cognition. *Annual Review of Psychology* 59(1): 617–645.

Beilock, Sian L. and Susan Goldin-Meadow 2010. Gesture changes thought by grounding it in action. *Psychological Science* 21(11): 1605–1610.

Bergen, Benjamin 2012. Louder Than Words: The New Science of How the Mind Makes Meaning. New York: Basic Books.

Bergen, Benjamin and Kathryn Wheeler 2010. Grammatical aspect and mental simulation. *Brain and Language* 112(3): 150–158.

Brunyé, Tad T., Tali Ditman, Carolyne R. Mahoney, Jason S. Augustyn and Holly A. Taylor 2009. When you and I share perspectives: Pronouns modulate perspective-taking during narrative comprehension. *Psychological Science* 20(1): 27–32.

Cienki, Alan and Cornelia Müller 2008. Metaphor, gesture, and thought. In: Raymond Gibbs (ed.), *The Cambridge Handbook of Metaphor and Thought*, 483–501. Cambridge: Cambridge University Press.

Chu, Mingyuan and Sotaro Kita 2011. The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General* 140(1): 102–115.

de Ruiter, Jan P. 2000. The production of gesture and speech. In: David McNeill (ed.), *Language and Gesture*, 284–311. Cambridge: Cambridge University Press.

Flusberg, Stephen and Lera Boroditsky 2011. Are things that are hard to physically move also hard to imagine moving? *Psychonomic Bulletin and Review* 18(1): 158–164.

Gibbs, Raymond 2006. Metaphor interpretation as embodied simulation. *Mind and Language* 21(3): 434–458.

Glenberg, Arthur M. 2010. Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science* 1(4): 586–596.

Goldin-Meadow, Susan, Susan L. Levine, Elena Zinchenko, Terina KuangYi Yip, Naureen Hemani and Laiah Factor 2012 Doing gesture promotes learning a mental transformation task better than seeing gesture. *Developmental Science* 15(6): 876–884.

Goodwin, Charles 2000 Action and embodiment within situated human interaction. *Journal of Pragmatics* 32(10): 1489–1522.

Hostetter, Autumn B. and Martha W. Alibali 2008. Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin and Review* 15(3): 495–514.

Hostetter, Autumn B. and Martha W. Alibali 2010. Language, gesture, action! A test of the Gesture as Simulated Action framework. *Journal of Memory and Language* 63(2): 245–257.

Kendon, Adam 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kita, Sotaro 2000. How representational gestures help speaking. In: David McNeill (ed.), *Language and Gesture*, 162–185. Cambridge: Cambridge University Press.

Kita, Sotaro and Aslı Özyürek 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48(1): 16–32.

Lakoff, George and Mark Johnson 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.

Levelt, Willem J. 1993. *Speaking: From Intention to Articulation*, Volume 1. Cambridge: Massachusetts Institute of Technology Press.

Louwerse, Max M. and Louise Connell 2011. A taste of words: Linguistic context and perceptual simulation predict the modality of words. *Cognitive Science* 35(2): 381−398.

Madden, Carol J. and Rolf A. Zwaan 2003. How does verb aspect constrain event representations? *Memory and Cognition* 31(5): 663−672.

Marghetis, Tyler and Rafael Núñez 2013. The motion behind the symbols: A vital role for dynamism in the conceptualization of limits and continuity in expert mathematics. *Topics in Cognitive Science* 5(2): 299−316.

Merleau-Ponty, Maurice 2002. *Phenomenology of Perception*. New York: Routledge.

McNeill, David 2005. *Gesture and Thought*. Chicago: University of Chicago Press.

Müller, Cornelia 2009. Gesture and Language. In: K. Malmkjaer (ed.), *Routledge's Linguistics Encyclopedia* (214−217). Abington/New York: Routledge.

Núñez, Rafael and Tyler Marghetis to appear. Cognitive linguistics and the concept(s) of number. In: Roi Cohen-Kadosh and Ann Dowker (eds.), *Oxford Handbook of Numerical Cognition*. Oxford University Press.

Parrill, Fey 2010. Viewpoint in speech-gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes* 25(5): 650−668.

Parrill, Fey 2011. The relation between the encoding of motion event information and viewpoint in English-accompanying gestures. *Gesture* 11(1): 61−80.

Parrill, Fey, Benjamin Bergen and Patricia Lichtenstein 2013. Grammatical aspect, gesture, and conceptualization: Using co-speech gesture to reveal event representations. *Cognitive Linguistics* 24(1): 135−158.

Pulvermüller, Friedemann and Luciano Fadiga 2010. Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience* 11(5): 351−360.

Rogers, William T. 1978. The contribution of kinesic illustrators toward the comprehension of verbal behavior within utterances. *Human Communication Research* 5(1): 54−62.

Sato, Manami 2010. Message in the body: Effects of simulation in sentence production. Unpublished Ph.D. dissertation, University of Hawai'i at Manoa.

Sassenberg, Uta and Elke Van Der Meer 2010. Do we really gesture more when it is more difficult? *Cognitive Science* 34(4): 643−664.

Stanfield, Robert A. and Rolf. A. Zwaan 2001. The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science* 12(2): 153−156.

Streeck, Jurgen 2008. Depicting by gesture. *Gesture* 8(3): 285−301.

Wagner Cook, Susan and Michael Tanenhaus 2009. Embodied communication: Speakers' gestures affect listeners' actions. *Cognition* 113(1): 98−104.

Winter, Bodo and Benjamin Bergen 2012. Language comprehenders represent object distance both visually and auditorily. *Language and Cognition* 4(1): 1−16.

Wu, Ying Choon and Seana Coulson 2007. How iconic gestures enhance communication: An ERP study. *Brain and Language* 101(3): 234−245.

Tyler Marghetis, San Diego (USA)
Benjamin K. Bergen, San Diego (USA)