case that embodied simulation is also involved in gesture *comprehension*.

The Gesture as Simulated Action (GSA) framework (Hostetter & Alibali 2008; 2010) explains how representational gestures might arise from an embodied cognitive system. Aligned with theories of embodied cognition (e.g., Barsalou 1999; Glenberg 1997), the GSA framework holds that language processing activates perceptual and motor states. Speakers form mental images of information they are expressing; these mental images activate the same perceptual and motor systems that are involved in perceiving and interacting with physical objects in the world. This activation of the motor system is sometimes overtly expressed as representational gestures.

Here, we consider whether simulated actions could also be involved in gesture *comprehension*. Niedenthal et al. argue that when people perceive smiles, they engage their motor systems, either to mimic the perceived smiles or to simulate the experience of performing the perceived smiles. This mimicry or simulation contributes to perceivers' understanding of the perceived smiles.

We suggest that the same mechanisms – mimicry and simulation – might also be involved in comprehending representational gestures. Imagine a speaker describing how he makes pizza, who uses gestures to depict how he uses his fingertips and his knuckles to flatten the dough. Imagine further that the speaker's addressee has never made pizza crust before. In comprehending the speaker's description and gleaning information from his gestures, the addressee might overtly mimic the speaker's actions, or she might mentally simulate such actions.

Do speakers overtly mimic other speakers' gestures? A growing body of evidence indicates that they do (e.g., Kimbara 2006; 2008; McNeill, in press). Mimicry does not arise as a simple coincidence of people using similar gestures when they talk about similar things. Instead, similarity of gesture form is more likely when interlocutors can see one another than when they cannot, suggesting that mimicry is purposeful (Kimbara, 2008). McNeill (in press) observes that researchers who code gestures often use mimicry as part of their effort to make sense of a speaker's gestures. He argues that addressees use mimicry to understand gestures that may not be readily interpretable otherwise.

Although overt mimicry of gestures does occur, it is not as widespread as might be expected if it were the primary mechanism of gesture comprehension. Recall that Niedenthal et al. emphasize that smile perceivers sometimes *simulate* the experience of the perceived smile, without overtly mimicking it. Along similar lines, it seems likely that gesture comprehenders sometimes simulate the gestures they perceive, without overtly producing them. These embodied simulations of gestural actions may contribute to comprehenders' interpretation of the perceived gestures. Support for this mechanism comes from studies that have shown activation of premotor areas when people observe actions (for reviews, see Jeannerod 2001; Rizzolatti et al. 2001). Premotor cortex is also activated when people observe gestures (e.g., Montgomery et al. 2007), and this activation is modulated by semantic information from the accompanying speech (Willems et al. 2007). It has been argued that, when addressees perceive gestures, they interpret them, at least in part, by activating the cortical networks involved in producing gestures, via an "observation-execution matching system" (e.g., Holle et al. 2008; see Dick et al. 2009, for discussion). This idea is similar to the one we advance here.

The SIMS model holds that, in the case of smiles, eye contact triggers embodied simulation. It is possible that eye contact may also initiate embodied simulation of the co-occurring gestures; however, no studies to date have specifically examined the relation between eye contact and gesture comprehension. Instead, the little research that has explored the role of eye gaze in gesture comprehension suggests that addressees are most likely to attend to speakers' gestures when speakers direct their gaze to the *gestures* (e.g., Gullberg & Holmqvist 1999). Further, addressees are more likely to incorporate information from the gesture into their comprehension when speakers have fixated on the gestures (Gullberg & Kita 2009).

One explanation for these findings is that speakers' gaze to their own gestures signals that the gestures are important; addressees then directly attend to those gestures and run "as-if" simulations of the gestures that facilitate comprehension. Speakers' shift in eye gaze may be particularly important for initiating addressees' simulations of speech-accompanying gestures, because addressees typically focus on speakers' faces, making it difficult to attend to detailed features of gestures produced in peripheral space.

Could mimicry and simulation account for individual and situational differences in addressees' reliance on gesture for comprehension? Research has shown that gestures make a greater contribution to comprehension when the material to be comprehended is complex, ambiguous, or challenging relative to the addressee's skills (Graham & Heywood 1976; McNeil et al. 2000). It also seems likely that individuals differ in their reliance on gesture for comprehension, although there is little data directly addressing this point. We suggest that addressees vary in how likely they are to engage in overt mimicry and/or simulation. Across contexts, individuals may be more likely to engage in mimicry and/or simulation in situations where language comprehension is more difficult, as suggested by McNeill (in press).

In sum, building on our past work about simulated action as an explanation for why speakers *produce* gestures, we suggest that similar mechanisms may be involved in *comprehension* of gestures. Viewing a speaker's overt simulation of an action or event may elicit overt mimicry, or it may evoke a corresponding simulation in the addressee's mind. In the SIMS model, mimicry and simulation contribute to perceivers' understanding of smiles. We suggest that mimicry and simulation also contribute to addressees' comprehension of speakers' gestures.

Emotion simulation and expression understanding: A case for time

doi:10.1017/S0140525X10001457

Marian Stewart Bartlett
Institute for Neural Computation, University of California, San Diego, 92093-0440.
marni@salk.edu **http://mplab.ucsd.edu/~marni**

Abstract: Niedenthal et al. present a model for embodied emotion simulation and expression understanding that spans multiple brain systems. This commentary addresses the potential role of time in this model, and its implications for understanding social dysfunction.

Niedenthal et al. diverge from an increasing trend for localizationalism in cognitive neuroscience, and present a model of emotional expression and understanding that spans the somatosensory system, motor system, reward system, amydgdala, basal ganglia, and prefrontal cortex, as well as systems involved in eye gaze and body posture. This model enfolds ideas relating to mirror neurons and the link between perception and production, but as a subset of a larger system for understanding emotions in others.

Not only is this model of emotion unlocalized in space, it is also not localized in time. The systems involved must communicate and pass activation through time in order to fully engage this emotion system, akin to a closed-loop dynamic process. It may be crucial that the timing of the interactions of these structures be right. Disorders of temporal processing at any level of this system, whether in the perception or the production of smiles, or in passing activation between brain structures, could disrupt this system and hence disrupt emotion understanding.

Consideration of the temporal dimension of such a model may lead to a better understanding of social deficits, such as those in autism spectrum disorders as well as those reported in association with attention deficit disorder. Indeed, Gepner and Feron (2009) describe a theory of temporal processing deficits that may underlie a range of deficits in autism. Is this theory at odds with the mirror neuron hypothesis (e.g., Williams et al. 2001), that a disruption of the motor neuron response to the perception of movement in others underlies social processing dysfunction in autism? Probably not. Both may be components of a larger emotion understanding system that involves multiple structures and their interactions in time. In support of this perspective, Oberman et al. (2009) showed that individuals with autism do indeed show spontaneous mimicry of facial expressions, but that the response is delayed.

Such models inform the development of interventions to help people with social dysfunction. The Niedenthal et al. model, together with theories such as the one proposed by Gepner and Feron (2009), suggests that social processing interventions should tap multiple processes, not individually, but together and at the right temporal intervals. Recent technology for automatically recognizing and responding to facial expression, head pose, and eye gaze in real-time opens up new possibilities for intervention systems that link perception and production on timescales related to social responding (Bartlett & Whitehill, in press; Cockburn et al. 2008). Such technology contributes not only to clinical research, but also to the study of learning and plasticity in perception and production systems, and to understanding the cognitive neuroscience of emotion.

"Smile down the phone": Extending the effects of smiles to vocal social interactions

doi:10.1017/S0140525X10001469

Frédéric Basso[a] and Olivier Oullier[b]

[a]*University of Rennes 1, CREM CNRS UMR 6211, IGR-IAE de Rennes 1, 35000 Rennes, France;* [b]*Laboratoire de Psychologie Cognitive (UMR 6146), Université de Provence & CNRS, Aix-Marseille Université, 13331 Marseille Cedex3, France.*

frederic.basso@univ-rennes1.fr olivier@oullier.fr
http://www.igr.univ-rennes1.fr/personnes/detail_fr_174.htm
http://www.oullier.fr

Abstract: The SIMS model offers an embodied perspective to cognition and behaviour that can be applied to organizational studies. This model enriches behavioural and brain research conducted by social scientists on *emotional work (*also known as *emotional labour)* by including the key role played by body-related aspects in interpersonal exchanges. Nevertheless, one could also study a more vocal aspect to smiling as illustrated by the development of "smile down the phone" strategies in organizations. We propose to gather face-to-face and voice-to-voice interactions in an embodied perspective taking into account Lakoff and Johnson's (1980) theory of conceptual metaphors.

Emotional work (also known as *emotional labour*) has been originally defined by Hochschild (1979, p.561) as "the act of trying to change in degree or quality an emotion or a feeling." It is noteworthy that research in that perspective started some thirty years ago, studying smiles of flight attendants, in a face-to-face social setting. But, following Hochschild's early suggestion, one should also consider how the effects of smiles can be effective during phone conversations, in a voice-to-voice setting when no visual information is shared (e.g., Sutton 1991).

Thanks to a substantial amount of field studies, data along that approach became available to support the presence of emotional work within organizations (Fineman 2000), whether for profit (e.g., between a sales representative and a customer; Ashforth & Humphrey 1993; Rafaeli 1989) or not (e.g., between a nurse and a patient; Froggatt 1998). By focusing on interactions and bodily cues such as facial expressions, eye contact, posture, and gestures, the SIMS model encapsulates the different aspects (and roles) a smile can play in all those contexts.

Nevertheless, the visual modality is not the only one through which smiles can be expressed. There is indeed a more vocal aspect to smiling as illustrated by the development of "smile down the phone" strategies in organizations (e.g., call centres). Because of the absence of face-to-face interactions during phone conversations, those "vocal smiles," including the tone of voice, constitute one of the keys to understanding emotional work in call centres (Belt et al. 2002; Taylor & Bain 1999). These centres are the illustration of the marketing logic known as *Customer Relationship Management*, aiming at developing long-term relationships between companies and their customers (Gans et al. 2003).

This managerial framework clearly constitutes an expression of emotional work centered around smiles down the phone to create empathy between the client and the sales representative (Richardson & Howcroft 2006). It appears that this vocal aspect to smiling is not considered in Niedenthal et al.'s SIMS. For instance, they focus on face-to-face interactions and do not take into account, so far, how efficient smiles can be in voice-to-voice ones.

Previous studies on emotional work complement the perspective offered by SIMS, especially concerning affiliative smiles (see sects. 2.2 and 6.1.1 of the target article). Hence, we suggest gathering face-to-face and voice-to-voice interactions in an embodied perspective – that is, still in opposition to an "amodal" view of knowledge (Lakoff & Johnson 1999). In our view, Barsalou's (1999) simulator would be enhanced by Lakoff and Johnson's theory of conceptual metaphors for it would ground SIMS in natural language – although this theory is not limited to the study of words.

Emotional work sheds new light on the metaphors used by actors in organizations. According to Froggatt (1998, p.332), the metaphorical language employed by nurses (e.g., draining and burden) reflects the emotional aspect of their work. Along the same line, Rees et al. (2007) argue for the existence of metaphors in describing the relationship between patients and physicians. Some therefore consider the emotional work of the physician towards the patient as a "metaphoric framework of clinical empathy" (Larson & Yao 2005, p. 1104). Here too, the challenge of creating empathy with others can be satisfied by using metaphors given that empathic processes at stake are at the core of the emotional work of the physician (Larson & Yao 2005).

Hence, metaphors are a way to access the emotional work of the actors in that the expression of emotions is metaphorical by essence. In line with Lakoff and Johnson (1980), Hochschild (2005, p. 344) stresses how metaphors "guide feeling, and, of course, feelings also guide metaphors." Similarly, when addressing emotional concepts, Lakoff (1987, p. 377) points that: "When we act on our emotions, we act not only on the basis of feeling but also on the basis of that understanding. Emotional concepts are thus very clear examples of concepts that are abstract and yet have an obvious basis in bodily experience."

Overall, it is noteworthy in this context that the three key elements used in SIMS to decrypt what lies beneath smiling face-to-face (perceptual cues, experiential cues, and conceptual knowledge; see sect. 6) can be found in voice-to-voice contexts.