European Journal of Social Psychology
Eur. J. Soc. Psychol. **39**, 1169–1172 (2009)
Published online in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/ejsp.678

Commentary

Prediction and emotion in dialogue

ARTHUR M. GLENBERG[1,2]*
[1]*Arizona State University, USA*
[2]*University of Wisconsin, USA*

One of the promises of the embodiment approach is the possibility of providing a unified account of psychology. As long as there is life, there is an acting body. Consequently, the body potentially shapes all psychological processes. The role of the body has been documented in cognition (Barsalou, 2008), social interaction (Semin & Smith, 2008), development (Thelen, 1995), education (Glenberg, 2008a), emotions (Niedenthal, 2007), and illness (Lindeman & Abramson, 2008). Pickering and Garrod (this volume) have provided an important contribution to the idea that the body shapes thought by developing an embodied approach to language that emphasizes its social nature. The primary goal of this commentary is to build on Pickering and Garrod by discussing some of the neuroscience of language, using that to provide a hypothesis regarding imitative and complementary activation of meaning, and then extending the Pickering and Garrod analysis into the domain of emotions.

As alluded to by Pickering and Garrod, it is quite likely that the human mirror neuron system (MNS) plays an important role in language (see also, Gallese & Lakoff, 2005; Glenberg, Sato, Cattaneo, Riggio, Palumbo, & Buccini, 2008; Rizzolatti & Craighero, 2007). Mirror neurons are found in Broca's area, other parts of premotor cortex, and parts of the parietal and temporal lobes. The intriguing property of mirror neurons is that they are equally strongly active when an animal takes a particular action (e.g., eating a peach) and when the animal sees another engage in the same action. If that animal is a human, then some mirror neurons will respond during action, observation of action, and when the person hears a description of the action (e.g., ''eat the peach,'' Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006; Tettamanti et al., 2005). Thus, the MNS provides a neural substrate for understanding the goals and intentions of others: I understand your action because resonance in my MNS provides the goal that I would have when taking the same action. That is, mirror neurons are the ultimate social information processors.

The MNS contributes to all levels of language, and hence demonstrates the importance of social interaction at all levels of language processing. For example, as noted by Pickering and Garrod, the MNS resonates upon hearing speech (Fadiga, Craighero, Buccino, & Rizzolatti, 2002), and it is likely that this activation drives changes in both speech perception and production (Goldinger, 1998; Guenther, Ghosh, & Tourville, 2006). The MNS also contributes to the meaning of the linguistic message (Aziz-Zadeh et al., 2006; Tettamanti et al., 2005). Most recently, there have been several analyses of how the MNS can contribute to the learning (Glenberg & Gallese, in preparation) and representation (Kemmerer & Castillo, in press) of syntax.

The MNS provides a neurophysiological mechanism for several of Pickering and Garrod's central themes. Because mirror neurons respond to both perception and action (including speech production), they demonstrate the use of a common representation in language perception and production. Furthermore, the MNS provides a mechanism for emulation and prediction: Motor resonance provides automatic access to the perceiver's goals and intentions which can be used to predict the producer's goals and intentions.

A model of the relations among the MNS, language, and prediction was sketched in Glenberg (2008b) and is developed more fully in Glenberg and Gallese (in preparation). As shown in Figure 1, paired controller and predictor models (e.g., for

*Correspondence to: Arthur M. Glenberg, Department of Psychology, Arizona State University, Tempe, AZ 85287-1104, USA.
E-mail: glenberg@asu.edu

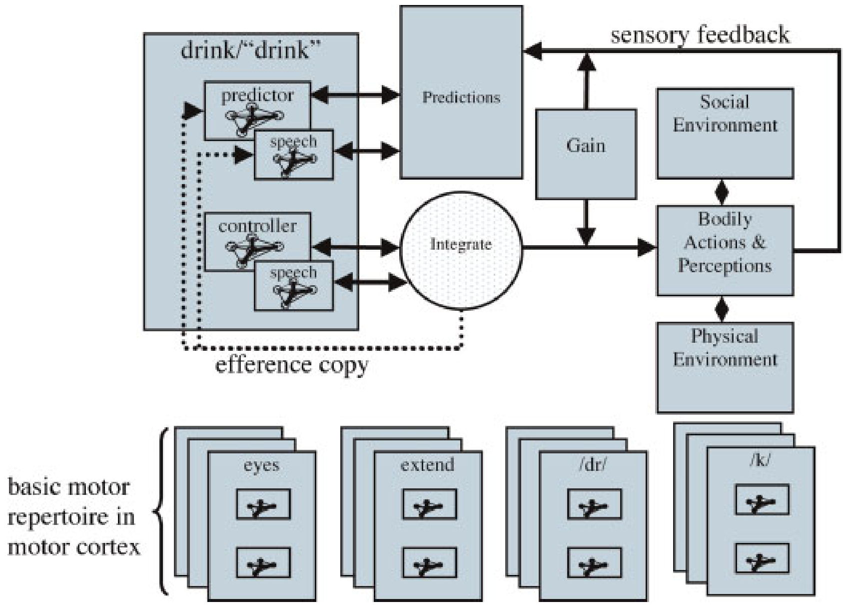

Figure 1. The Action-based Language (ABL) theory (Glenberg & Gallese, in preparation) is an adaptation of Wolpert and colleagues (e.g., Wolpert, Doya, & Kawato, 2003) HMOSAIC model of action control. Two levels of hierarchy are illustrated. At the bottom of the figure are modules controlling simple motor acts such as moving the eyes or pronouncing a phoneme. The main part of the figure illustrates a higher-level module for controlling a more complex act such as drinking, as well as the pronunciation of ''drink.'' Each module consists of two types of mirror-neuron neural networks. The controller sends commands to the muscles (or at higher levels, orders the selection of simple motor acts). Often, several related commands must be integrated, and the integrated command is fed back to the predictor neural networks. The predictor anticipates sensory feedback as well as changes in the physical and social environment as a result of action, and thus is equivalent to Pickering and Garrod's emulator. Gain control dampens literal action when using the module for planning and language comprehension

the action of drinking) provide hierarchical control over the basic motor acts that, when structured appropriately, produce complex actions such as drinking. The neural networks that comprise the controllers and predictors are envisioned as consisting, at least in part, of mirror neurons.

The controller generates action plans, and a copy of these commands (the efference copy) is fed back to the predictor neural network. Predictions derived from the efference copy can be of changes in sensation (e.g, I will feel the water as I drink), changes in the environment (e.g., the level of water in the glass will decrease), and changes in the social environment (e.g., I will hear requests from others who wish to drink). Thus the predictor model captures Pickering and Garrod's notion of emulation.

An important feature of the Glenberg and Gallese scheme is that there is a speech controller (e.g., how to pronounce the word ''drink'') closely associated with the action controller. It is this association that grounds the meaning of the word. For example, the meaning of the verb ''to drink'' consists of the actions of drinking and the predictions that flow from those actions. Furthermore, the association between word pronunciation and action accounts for the close relation between speech and gesture (e.g., Hostetter & Alibali, 2008): Pronouncing the word activates the actions used in accomplishing goal.

Consider now, how such a system can produce imitative and complementary responses for both form and meaning. To extend Pickering and Garrod's example, suppose that Carol is relating to David an interaction between Alice and Bill. ''They were having a terrible argument when Alice lost control and slapped Bill's face.'' At the level of form, David's mirror neurons in his controller for pronouncing the word ''slap'' will resonate upon hearing the initial articulation of the word. This resonance will produce activation of the commands for pronouncing the word, that is, an imitative response. Depending on the level of gain control (see Figure 1), David may utter the word, move his lips, or simply use the efference copy to generate predictions about Carol's pronunciation of the word.

Activation of the neural network for pronouncing ''slap'' will also activate David's network for producing the slapping action. This activation may be enhanced by Carol beginning to gesture slapping (by opening her hand and pulling her arm back). That is, David's mirror neurons in his slap action controller will resonate upon the perception of Carol's gesture. Activation of his action controller provides an explanation for the imitative response (e.g., David making the gesture) at the level of meaning. Furthermore, the efference copy of the slap commands is fed back to David's predictor neural network, thus providing a basis for complementary responses such as David predicting Bill's behavior.

As illustrated by the Alice and Bill vignette, language is often used to convey information about emotions, and there is some work investigating the embodiment of the interaction between language and emotion. For example, Mondillon and Niedenthal (as reported in Niedenthal, 2007) demonstrated using EMG recordings that upon processing the meaning of words conveying emotions, there is selective activation of the facial musculature. For example, positive words activate the zygomaticus muscle that is used in smiling, and negative words activate the corrugator muscle used in furring the brow.

Havas, Glenberg, and Rinck (2007) traced the effect of the embodiment of emotions on sentence understanding. In their experiments, a participant held a pen in the mouth (so that the pen pointed out of the mouth) using only the teeth (which produces a smile) or only the lips (which produces a pout and prevents smiling). This technique reliably brightens or dampens affect (e.g., Strack, Martin, & Stepper, 1988). While holding the pen, the participants read sentences describing pleasant and unpleasant events and judged them as easy or hard to understand. The experiment revealed an interaction between the way the pen was held and the emotional content of the sentence: Participants were faster to judge pleasant sentences when smiling and faster to judge negative sentences when smiling was prevented. Thus, just as understanding the word ''slap'' is facilitated by activating the motor program for slapping, understanding the emotional content of the language depended, at least in part, on being able to simulate or produce the corresponding emotion.

Havas, Glenberg, Gutowski, Lucarelli, and Davidson (in preparation) used an unusual manipulation to demonstrate the effect of corrugator activation on language understanding. In an initial experiment, Havas et al. (in preparation) replicated the Mondillon and Niedenthal study, but using sentences. Sentences with positive emotional content activated the zygomaticus muscle, and sentences with negative content (sadness and anger) activated the corrugator muscle. In a second experiment, participants read the same sentences divided into two sets of happy, sad, and angry sentences. The first set was read shortly before the participant received a treatment of cosmetic botulinim toxin in the corrugator muscle. The cosmetic use of the toxin paralyzes the muscle, which has the effect of reducing glabellar (frown) lines in the forehead. The second set of sentences was read 2 weeks later, near to when the toxin reaches its maximum effect and the patient cannot produce frown lines using the corrugator muscle. As expected, the toxin had no effect on the reading time of happy sentences, because the corrugator muscle is not activated during the expression of happiness. In contrast, reading of the angry sentences was slowed after the injection. That is, to the extent that the participants could no longer simulate (or produce) anger, they were slowed in understanding the angry sentences.

Consider how this embodiment of emotion, along with the mechanisms illustrated in Figure 1, leads to imitative and complementary responding regarding the emotional content of a conversation. Suppose again that Carol is recounting the Alice and Bill vignette to David. If Carol is frowning or looking concerned, Bill's MNS will bias him to frown. Of course, there is much data illustrating chameleon effects (Barsalou, Niedenthal, Barbey, & Rupert, 2003; Chartrand & Bargh, 1999) such as this. Now, with David frowning, we know from Havas et al. (2007, in preparation), that he will more readily process the negative content of the vignette. Thus, David's frown is an imitative response that facilitates understanding of the sentence and production of complementary responses such as inferring Bill's state. Suppose, however, that Carol is relating the vignette with a smile. In this case, David's imitative smiling will slow his understanding of the content of the vignette. However, his imitative smiling will ultimately help him to understand Carol's expression of schadenfreude.

In summary, the mechanism of prediction during dialogue makes multiple contributions to understanding. As Pickering and Garrod suggest, prediction facilitates the processing of language in dialog. Furthermore, as extended here, prediction facilitates the exchange of emotional information important both for understanding the message behind the words and the smooth functioning of social interactions.

ACKNOWLEDGEMENTS

Preparation of this commentary was facilitated by a grant from the National Science Foundation (BCS 0744105) to Arthur Glenberg. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Embodied semantics and the premotor cortex: Congruent representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, *16*, 1818–1823.

Barsalou, L. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.

Barsalou, L. W., Niedenthal, P. M., Barbey, A., & Ruppert, J. (2003). Social embodiment. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 43–92). San Diego: Academic Press.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*, 399–402.

Gallese, V., & Lakoff, G. (2005). The Brain's concepts: The role of the sensory-motor system in reason and language. *Cognitive Neuropsychology*, *22*, 455–479.

Glenberg, A. M. (2008a). Embodiment for education. In P. Calvo, & A. Gomila (Eds.), *Handbook of cognitive science: An embodied approach* (pp. 355–372). Amsterdam: Elsevier.

Glenberg, A. M. (2008b). Toward the integration of bodily states, language, and action. In G. Semin, & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 43–70). Cambridge: Cambridge University Press.

Glenberg, A. M., & Gallese, V. (in preparation). Action-based language.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology*, *61*, 905–919.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251–279.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain & Language*, *96*, 280–301.

Havas, D., Glenberg, A. M., Gutowski, K., Lucarelli, M., & Davidson, R. (in preparation). Cosmetic use of botulinim toxin affects processing of emotional language.

Havas, D. A., Glenberg, A. M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review*, *14*, 436–441.

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

Kemmerer, D. M., & Castillo, J. G. (in press). The two-level theory of verb meaning: An approach to integrating the semantics of action with the mirror neuron system. *Brain and Language*.

Lindeman, L., & Abramson, L. (2008). The mental simulation of motor incapacity in depression. *Journal of Cognitive Psychotherapy*, *22*, 228–249.

Niedenthal, P. M. (2007). Embodying emotion. *Science*, *316*, 1002–1005.

Rizzolatti, G., & Craighero, L. (2007). Language and mirror neurons. In G. Gaskell (Ed.), *Oxford handbook of psycholinguistics*. Oxford: Oxford University Press.

Semin, G., & Smith, E. R. (Eds.). (2008). *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches*. Cambridge: Cambridge University Press.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality & Social Psychology*, *54*, 768–777.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, *17*, 273–281.

Thelen, E. (1995). Motor development: A new synthesis. *American Psychologist*, *50*, 79–95.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London B*, *358*, 593–602.

DOI: 10.1002/ejsp