

Special issue article

Prediction and embodiment in dialogue

MARTIN J. PICKERING^{1*} AND SIMON GARROD²

¹*University of Edinburgh, UK*

²*University of Glasgow, UK*

Abstract

We argue that embodiment (via use of action-based representations) plays a crucial role in dialogue. To illustrate the argument we use studies of language comprehension. We first compare two distinct literatures, one concerned with the activation of non-linguistic action-based representations of meaning, and the other with representations of linguistic form associated with language production. We then argue that both types of embodiment support emulation and prediction. Hence, such embodiment enables addressees to anticipate both what their partner is likely to say next and what she is likely to do. We conclude by suggesting that such anticipation is essential for fluent and timely social interactions. Copyright © 2009 John Wiley & Sons, Ltd.

Most research in the adult psychology of language attempts to explicate the mechanisms of language processing as an end in itself, rather than to determine the cognitive underpinnings of the social activity of dialogue. This means that psychologists concerned with social interaction typically pay little attention to existing research in psycholinguistics. However, an understanding of processing mechanisms may help explain why people are remarkably adept at dialogue given its cognitive complexity. In this paper, we argue that such mechanisms are embodied and that such embodiment facilitates dialogue.

Our claim that language processing is embodied involves two components. The first is that processing involves the use of representations associated with the actions that the language refers to. For example, producing or comprehending *walk* involves representations involved in the action of walking. Since walking involves motor programs that lead to putting one foot in front of the other, the embodied account requires that some aspects of these programs are involved in processing the word *walk*. This use of embodied (or grounded) cognition corresponds to that of Barsalou (2008) and Glenberg (2008). We refer to this as embodiment of meaning.

The second sense in which language processing is embodied is that it involves representations associated with the actions of producing language. Clearly this is definitionally the case for language production, but is a much more interesting claim in the case of language comprehension. We refer to this as embodiment of form. The distinction between embodiment of form and meaning is echoed by Gallese's (2008) distinction between simulation at the vehicle level and simulation at the content level (see also Zwaan & Taylor, 2006).

We argue that embodiment of meaning and form is fundamental to the success of real-time, interactive language use—in other words, of dialogue. Many characteristics of dialogue suggest that it should be very hard (Garrod & Pickering, 2004). Interlocutors have to produce responses on the fly, to comprehend incomplete or elliptical utterances, to repeatedly determine who to speak to and precisely when to speak, to comprehend and produce at the same time, and of course to switch constantly between comprehension and production. Yet almost everyone (including young children) can hold a

*Correspondence to: Martin J. Pickering, Department of Psychology, University of Edinburgh, George Square, Edinburgh EH8 9JZ, UK.
E-mail: martin.pickering@ed.ac.uk

conversation (whereas giving a speech is highly challenging). For example, people constantly use it to negotiate social relations (e.g. in family discussions), to coordinate routine social activities (e.g. purchasing an item in a store), and to influence others (e.g. in informal debate). In fact, conversation is a paradigm example of real-time joint activity. It requires continuous management of interpersonal relations and needs to be analysed at an interpersonal level.

For joint activity to be successful, it is not generally sufficient for interactants to interpret each other's actions after they occur. Instead, they need to predict those actions and integrate the representations of those actions with their own (Sebanz, Bekkering, & Knoblich, 2006). Dialogue is no exception. Interlocutors activate the embodied representations associated with meaning and form in order to predict each other's linguistic and non-linguistic actions and to integrate those representations with their own.

In this paper, we focus on language comprehension. We demonstrate that it involves considerable activation of representations associated with action (or motoric representations). In line with our distinction between meaning and form, some of this evidence shows activation of representations associated with actions related to the utterance's meaning (meaning-based representations), and the other relates to representations associated with language production (form-based representations). We then discuss how such embodiment sometimes leads to overt imitation (mimicry) and sometimes leads to complementary responses. We next ask how embodiment supports prediction. There is good (but not conclusive) evidence that form-based representations serve to make predictions during comprehension in monologue and dialogue, and that they facilitate overt responding in dialogue. Here, we make the case that meaning-based representations serve a similarly predictive role, and that form and meaning representations become integrated across production and comprehension. We conclude by relating embodiment in dialogue to social interaction more generally.

EMBODIMENT OF MEANING

As is well known, there is a great deal of evidence for embodiment related to meaning of what is being said during language comprehension. For example, people's representations of utterances that describe scenes incorporate spatial perspective, which clearly relates to the way in which they behave in the actual world (e.g. Bransford & Johnson, 1973). More recent evidence indicates that embodied representations affect the process of comprehension and therefore must be constructed 'on-line'. Thus, Glenberg and Kaschak (2002) found that participants were faster at judging a sentence describing movement in one direction (e.g. *close the drawer*) as sensible if the judgement response involved movement in the same direction (in this case, away from the body). Similarly, Zwaan and Taylor (2006) had participants turn a knob to present sentences one word at a time. Participants were faster to read words (e.g. *turned down*) when they turned the knob in the direction implied by those words (in this case, anticlockwise).

These findings suggest rapid, functional activation of action-based schemas during language comprehension. This is supported by evidence showing similarly rapid activation of motor systems in the brain (Pulvermüller, 2005). For example, an MEG study showed activation of appropriate cortical motor areas after presentation of *potki* (kick) versus *hotki* (eat), within 170 ms of when they could be uniquely identified (Pulvermüller, Shtyrov, & Ilmoniemi, 2005). This speed of activation is comparable with the speed of word recognition and therefore strongly suggests that motor activation occurs 'on-line'.

EMBODIMENT OF FORM

Language comprehension activates action-based schemas associated with the form of utterances as well as with their meaning. Thus, there is direct evidence for involvement of articulation in speech comprehension, with listeners activating appropriate muscles in the tongue and lips while listening to speech but not non-speech (Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Watkins, Strafella, & Paus, 2003). Additionally, increased muscle activity in the lips is associated with increased activity (i.e. blood flow) in Broca's area, suggesting that this area mediates between the comprehension and production systems during speech perception (Watkins & Paus, 2004). Other fMRI studies have also demonstrated a large overlap between the cortical areas active during speech production and those active during passive listening (Pulvermüller,

Huss, Kherif, Moscoso del Prado Martin, Hauk, & Shtyrov, 2006; S.M. Wilson, Saygin, Sereno, & Iacobini, 2004). Thus, comprehension activates the production system, and leads to covert imitation. (Notice that the motor activation is likely to follow from the activation of the production system rather than itself be the cause of imitation.) There is also evidence for the activation of brain areas associated with production during aspects of comprehension from phonology (Heim, Opitz, Müller, & Friederici, 2003) to narrative structure (Mar, 2004). Some of this evidence is used to provide support for the motor theory of speech perception, according to which speech comprehension involves the direct perception of motor gestures (Galantucci, Fowler, & Turvey, 2006).

THE EFFECTS OF EMBODIMENT

To explain how embodiment occurring during language comprehension affects social interaction, we should consider its effects on (overt) behaviour, and specifically those behaviours that are likely to have social consequences. In fact, both embodiment of meaning and embodiment of form can promote two kinds of effects, one leading to overt imitation, and the other leading to complementary responses. To see this, consider the possible embodiment effects of hearing someone say *slap*. At the meaning level, overt imitation would lead to the act of slapping. However, a complementary response might lead to flinching (to avoid the slap). At the form level, overt imitation might lead to the comprehender uttering *slap*, whereas a complementary response might lead to the comprehender uttering *his face* (i.e. words that could occur next).

Notice that the imitative and complementary responses are closely related—they correspond to the (linguistic or non-linguistic) action and the immediate response to that action. This will become clear when we consider the process of emulation below. Thus, our distinction does not correspond to the distinction between assimilation and contrast effects in social cognition, in which some responses are assimilated to a prime stimulus (e.g. Higgins, Rholes, & Jones, 1977) and some are contrasted with it (e.g. Herr, Sherman, & Fazio, 1983; see Förster, Liberman, & Kuschel, 2008).

In relation to embodiment of form, we see very strong evidence for spontaneous overt imitation at many linguistic levels, such as choice of words (Brennan & Clark, 1996) and choice of grammar (Branigan, Pickering, & Cleland, 2000; see Pickering & Ferreira, 2008). This implies that people construct imitative plans at the relevant stages in the production process. Much of this evidence comes from studies of dialogue, but there is also evidence for linguistic imitation in the absence of an interlocutor (e.g. Bock, Dell, Chang, & Onishi, 2007). For imitation to be used by the comprehension system, it must occur very quickly and result from automatic processes rather than conscious decisions. In fact, phonological or acoustic imitation is extremely rapid (Fowler, Brown, Sabadini, & Weihsing, 2003) and the faster the imitation, the more faithful it is (Goldinger, 1998). Furthermore, interlocutors are almost entirely unaware that they imitate each other's grammar or choice of words (see Pickering & Garrod, 2004).

However, there is also complementary activation of form. Thus, interlocutors often complete each other's speech appropriately. Tannen (1989) considers such completions in terms of an analysis of a dinner-party conversation; for example, one interlocutor describes a pianist's hands as *Just completely soft and limp*, and another interlocutor utters the nearly synonymous *mush* at the same time as *limp*. Similarly, Clark and Wilkes-Gibbs (1986) note that speakers describing abstract shapes in a matching game sometimes pause during a description, in a way that allows their addressee to complete it (e.g. A: *And number 12 is, uh, . . .* B: *Chair*. A: *With the chair, right*. B: *Got it*). In addition, people are faster at naming a word when it is compatible with previous context than when it is not. For example, Tyler and Marslen-Wilson (1977) found that people were faster naming *Are* than *Is* following *As they glide gracefully over the city, flying kites*. In addition, Griffin and Bock (1998) found that participants were faster to name a picture after hearing a sentence context that strongly predicted that word than after hearing a sentence context that did not strongly predict any particular word.

Similarly, there is evidence for both imitative and complementary activation of meaning. We have already referred to both behavioural and neuroscientific evidence for the activation of imitative representations (Glenberg & Kaschak, 2002; Pulvermüller et al., 2005; Zwaan & Taylor, 2006). With respect to complementary activation, Glover, Rosenbaum, Graham and Dixon (2004) presented participants with words that referred to small or large objects (as part of an apparent memory task) and found that the size of the object affected hand aperture in a subsequent object grasping test. In other words, the word activated aspects of meaning relating to how people would act on the object, not aspects of meaning relating to the object itself. Additionally, Chambers, Tanenhaus, Eberhard, Filip, and Carlson (2002) gave participants instructions such as *Pick up the cube. Now put it inside the can*, in a context where there was one cube and two cans, only

one of which could accommodate the cube. Participants looked directly at the appropriate can. Since *the can* is ambiguous, the looks to the appropriate can were presumably the result of interpreting *the cube*.

We might now ask why both form and meaning embodiment sometimes lead to imitation and sometimes to complementarity. A functional explanation might stress that it is sometimes adaptive to imitate (e.g. to run when others run) and sometimes to behave in a complementary fashion (e.g. to flinch when the object of aggression), and clearly such an explanation can be carried over to language. In more cognitive terms, an appropriate explanation may involve mechanisms of suppression used to inhibit one's own responses (e.g. Dell, 1986), and any explanation of activation of form has to deal with a very detailed experimental literature (see Hartsuiker, Pickering, & De Jong, 2005, for discussion). We can ask similar questions about the effects of meaning embodiment (e.g. Kaschak & Borreggine, 2008).

EMBODIMENT AS EMULATION

We have suggested that such embodiment does not just constitute a behavioural response, but helps people to predict appropriate linguistic and non-linguistic responses to linguistic stimuli. In previous work, we have argued that people use the production system to make form-based predictions during language comprehension, in a way that assists the comprehension of monologue (Pickering & Garrod, 2007). Here we argue that people make form- and meaning-based predictions, and that they do so when using language in social interaction.

Specifically, we propose that embodiment of both meaning and form acts as an emulator (Grush, 2004). An emulator is a kind of forward model of an external system that runs simulations of that system in real time. For example, people appear to rapidly construct models of actual movements before making those movements and use those models as guides for corrections. Thus before moving your arm, you model the path it should take, and if it deviates from that path, you correct accordingly. Such an emulator can be used much more rapidly than real-world feedback, and additionally provides information about where your arm is likely to be, in the absence of accurate sensory information (see Desmurget & Grafton, 2000; Wolpert & Flanagan, 2001).

Pickering and Garrod (2007) proposed that the production system acts as an emulator during language comprehension. At each step the emulator predicts the next element in the speech stream and uses this prediction to aid understanding. One possibility is that this prediction simply facilitates the process of understanding. On this account, understanding is generally difficult (in particular difficult to do at the necessary speed), and prediction serves to make it more manageable (See Glenberg, 2008, for a similar proposal about action-based emulation). This proposal keeps understanding and prediction separate. However, a more radical alternative is that prediction is inherent to the notion of understanding itself. Thus, Prinz (2006) argued that understanding an action should encompass both identification of ongoing and anticipation of upcoming action.

If we apply this to language, it suggests that the process of understanding an utterance involves both interpreting what has actually been heard and anticipating what might come next, whether linguistic or non-linguistic. Importantly, this can mean either what the current speaker or the current addressee (i.e. the understander) is likely to say or do. For example, understanding *Run away!* would involve interpreting the words, but would also involve (1) considering a query such as *Why?*; (2) anticipating an explanation by the current speaker; (3) considering running away; (4) anticipating that the current speaker might start running away. In this case, (1) and (2) involve linguistic prediction, while (3) and (4) involve non-linguistic prediction; (1) and (3) involve predicting one's own actions, while (2) and (4) involve predicting the speaker's actions. It should be clear that the possibility of making such predictions follows from the emulation account mentioned above.

However, why should emulation be especially helpful to understanding? The argument is that one's interlocutor's linguistic system is highly similar to one's own, so emulating within one's own system is an excellent proxy for one's interlocutor. This argument holds for social interaction generally. M. Wilson and Knoblich (2005) argue that emulation serves as an important basis for perception of actions by conspecifics. You use your own motor programs to emulate the behaviour you are viewing in others. Some evidence comes from the finding that people are better predicting a movement trajectory (e.g. in dart-throwing or handwriting) when viewing a video of themselves versus others (see M. Wilson & Knoblich, 2005). Presumably, predicting with one's motor program is most accurate when the object of the prediction is one's own actions.

However, in the case of dialogue, the relevant aspects of interlocutors' mental states are considerably more similar to each other than is the case for generic conspecifics. Trivially, interlocutors must speak the same language, so that they have very similar representations of sounds, words, grammar, and the like. However, more than this, their mental states become aligned as a result of the interaction. Dialogue is successful to the extent that interlocutors come to see the world in similar ways to each other. Pickering and Garrod (2004) argue that such alignment does not involve extensive modelling of the each other's minds, but rather proceeds by largely unconscious priming of linguistic representations at many different levels. The effect of this is that interlocutors typically end up with similar patterns of activation for sounds, words, grammar and so on—in other words, aligned representations of form. In addition, they activate the same meaning representations as each other, so that these become aligned as well. This therefore means that emulation is likely to provide a particularly good proxy for the interlocutor at both the form and the meaning level, and hence acts of emulation are likely to be accurate in representing one's interlocutor's mental states. In other words, emulation should contribute to a linguistic 'shared reality'.

EMBODIMENT AND SOCIAL INTERACTION

It should therefore be clear how embodiment, combined with emulation, is likely to support social interaction. Emulation of meaning helps both to predict what and when your partner is going to do something and to predict (in the sense of forward planning) what you should do next. In real face-to-face conversations actions merge in a seamless way with the production and comprehension of the speech. For example, Clark and Krych (2004) manipulated whether a director could see a builder's workspace (and hence whether the space was shared) during a LEGO construction task. When the workspace was visible the director used many deictic expressions (e.g. *here, there, this, that, like this*) and timed the speech to fit in with the builder's sequence of actions. She also provided many demonstrations (deliberately positioning objects in ways to elicit responses from the builder). It appeared that the builder's actions were treated as continuous feedback by the director and directly affected both the content and timing of the speech.

Embodiment of form should also help social interaction. An interesting example of the latter is with the timing of turn-taking during conversation. There is good evidence that listeners during conversation predict precisely when the current speaker's turn will end and when they should begin speaking (de Ruiter, Mitterer, & Enfield, 2006). Models of inter-turn intervals suggest that speakers and listeners synchronise their syllabic speech rates during dialogue for this purpose (M. Wilson & Wilson, 2005). This process involves entrainment of the speech production system by the comprehension system. In other words, your speech production system emulates your comprehension system to predict when your interlocutor's turn should finish and when you should start your turn.

So far, we have focused on embodiment relating to the actual linguistic forms that are encountered. However, of course, there are many effects of these forms on other aspects of processing. Such effects are often considered under the general heading of social priming. For example, Bargh, Chen, and Burrows (1996) presented people with words such as *Florida* and found that this caused them to walk slowly. This can be seen as an embodied effect of language comprehension. The important difference from the studies considered above (e.g. Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006) is that this study involved an association between concepts. In this case, words associated with the elderly primed a concept of the elderly, and this concept primed slowness. (Of course, some such effects involve contrast rather than assimilation; e.g. Dijksterhuis, Spears, & Lépinasse, 2001). Much of the literature on social priming uses linguistic stimuli, and the process of comprehending those stimuli and activating embodied representations can lead to socially relevant behaviour.

More generally, recent research has addressed the question of the cognitive and neural processes that underlie the construction and interpretation of social signals (e.g. Frith & Frith, 2007). In particular, this research helps us understand why people can be successful in many forms of joint action (e.g. Sebanz et al., 2006), and how social-cognitive processes can underlie interaction (see Semin & Cacioppo, 2008). In such explanations, it is critical that people are able to construct common representations of their own and each other's actions. Whereas the research provides theories of how such representations can be developed in perceptual-motor processes (e.g. Hommel, Müsseler, Aschersleben, & Prinz, 2001), this approach has not been applied to the underpinnings of linguistic communication.

One effect of this perspective for both social and cognitive psychology is that it provides an alternative to a traditional 'sandwich' account, in which amodal cognitive processes constitute a set of central mechanisms that are isolated from input and output processes concerned with perception and action, respectively (e.g. Fodor, 1983; see Mahon & Caramazza, 2008). Instead, perception involves the immediate activation of action-based representations concerned with form or

meaning or both. In social psychological terms, this is straightforwardly consistent with the perception-behaviour expressway account (Dijksterhuis & Bargh, 2001). For the psycholinguist, this leads to a focus on interactive language use rather than decontextualised language (typically involving isolated participants responding to stimuli in booths). If we acknowledge that the primary site for language use is interactive dialogue and that much language comprehension (at least) involves the activation of embodied representations, the psychology of language can be seen as the socially grounded study of the mechanisms of communication, and therefore an excellent bridge between social and cognitive psychology.

CONCLUSIONS

We have argued that embodiment plays a crucial role in conversation. Embodiment both in relation to the meaning of what is being said and its form supports emulation and prediction. Hence, it enables communicators to anticipate both what their partner is likely to say next and what she is likely to do. We suggest that such anticipation is essential for fluent and timely social interactions. In conclusion, our account helps explain why conversation is so easy, and shows how the processes that take place during embodied comprehension of language facilitate social interaction in general.

ACKNOWLEDGEMENTS

The authors acknowledge the support of ESRC Grant no. RES-062-23-0376.

REFERENCES

- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype-activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.
- Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*, 437–458.
- Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic coordination in dialogue. *Cognition*, *75*, B13–B25.
- Bransford, J. D., & Johnson, M. K. (1973). Consideration of some problems of comprehension. In W. G. Chase (Ed.), *Visual Information Processing*. New York: Academic Press.
- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1482–1493.
- Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language*, *47*, 30–49.
- Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*, 62–81.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.
- De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, *82*, 515–535.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.
- Desmurget, M., & Grafton, S. (2000). Forward modelling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences*, *4*, 423–431.
- Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. In M. P. Zanna (Ed.), *Advances in experimental social Psychology* (Vol. 33, pp. 1–40). San Diego: Academic Press.
- Dijksterhuis, A., Spears, R., & Lépinasse, V. (2001). Reflecting and deflecting stereotypes: Assimilation and contrast in impression formation and automatic behavior. *Journal of Experimental Social Psychology*, *37*, 286–299.
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, C. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*, 399–402.
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Förster, J., Liberman, N., & Kuschel, S. (2008). The effect of global versus local processing styles on assimilation versus contrast in social judgment. *Journal of Personality and Social Psychology*, *94*, 579–599.
- Fowler, C. A., Brown, J. M., Sabadini, L., & Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language*, *49*, 396–413.
- Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology*, *17*, R724–R732.

- Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*, 361–377.
- Gallese, V. (2008). Mirror neurons and the social nature of language: The neural exploitation hypothesis. *Social Neuroscience*, *3*, 317–333.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*, 8–11.
- Glenberg, A. M. (2008). 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., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*, 558–565.
- Glover, S., Rosenbaum, D. A., Graham, J., & Dixon, P. (2004). Grasping the meaning of words. *Experimental Brain Research*, *154*, 103–108.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251–279.
- Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, *38*, 313–338.
- Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, *27*, 377–435.
- Hartsuiker, R. J., Pickering, M. J., & De Jong, N. (2005). Semantic and phonological context effects in speech error repair. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 921–932.
- Heim, S., Opitz, B., Müller, K., & Friederici, A. D. (2003). Phonological processing during language production: fMRI evidence for a shared production-comprehension network. *Cognitive Brain Research*, *12*, 285–296.
- Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, *19*, 323–340.
- Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*, 141–154.
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–878.
- Kaschak, M. P., & Borreggine, K. L. (2008). Temporal dynamics of the action-sentence compatibility effect. *The Quarterly Journal of Experimental Psychology*, *61*, 883–895.
- Mahon, B., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology—Paris*, *102*, 59–70.
- Mar, R. A. (2004). The neuropsychology of narrative: Story comprehension, story production and their interrelation. *Neuropsychologia*, *42*, 1414–1434.
- Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*, 427–459.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169–225.
- Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*, 105–110.
- Prinz, W. (2006). What re-enactment earns us. *Cortex*, *42*, 515–517.
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews. Neuroscience*, *6*, 576–582.
- Pulvermüller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, *103*, 7865–7870.
- Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience*, *17*, 884–892.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*, 70–76.
- Semin, G. R., & Cacioppo, J. T. (2008). Grounding social cognition: Synchronization, entrainment, and coordination. In G. Semin, & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 119–148). Cambridge: Cambridge University Press.
- Tannen, D. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.
- Tyler, L. K., & Marslen-Wilson, W. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 683–692.
- Watkins, K. E., & Paus, T. (2004). Modulation of motor excitability during speech perception: The role of Broca's area. *Journal of Cognitive Neuroscience*, *16*, 978–987.
- Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *41*, 989–994.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*, 460–473.
- Wilson, S. M., Saygun, A. P., Sereno, M. I., & Iacobini, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*, 701–702.
- Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review*, *12*, 957–968.
- Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*, 729–732.
- Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*, 1–11.