

Chapter 9

Experiential traces and mental simulations in language comprehension

Rolf A Zwaan

9.1 Experiential traces

The goal of this chapter is to put forth a framework for an embodied, experiential view of language comprehension. The claim is not that the elements of this framework are entirely new. The contribution lies more in the integration from insights from the diverse literatures on concepts, episodic and semantic memory, psycholinguistics, computational linguistics, motor control, and action observation. The key notion in the framework is that of the experiential trace. Experimental traces are laid down in multiple modality-specific cortical areas (e.g., Barsalou 1999; Damasio 1999). As such, they are multimodal records of experience. Importantly, linguistic constructions such as words and standard phrases are also multimodal experiential traces (Sadoski and Paivio 2001). For ease of exposition, a distinction will be made between linguistic traces and referential traces, but the key is that the two types of traces are not intrinsically different from one another (Zwaan and Madden 2005).

Associations between co-occurring linguistic traces and referential traces, *L–R associations*, are formed via Hebbian learning. Often, such co-occurrences emerge during *embedded comprehension* (Spivey and Richardson, in press), in which language is used to refer to the immediate environmental context and reference is established via joint attention (e.g., eye gaze, pointing). Developmental research has established that in such contexts young language learners use certain heuristics in resolving reference. For example, they tend to associate new linguistic constructions with whole objects rather than with their parts. They also associate new linguistic constructions with new entities, rather than with items for which they already have a lexical associate.

An additional way in which L–R associations can be formed is via *referential bootstrapping*. For example, a website for elementary school students features the following riddle: “What has horns like a giraffe, a deer head, a horse neck, and legs like a zebra?” The answer is, of course, an okapi. Although it may not be possible to construct a completely accurate visual representation of an okapi based on the riddle, this representation will probably be sufficiently accurate to enable one to recognize an okapi upon being exposed to this rare mammal, either in the flesh or in the form of a picture.

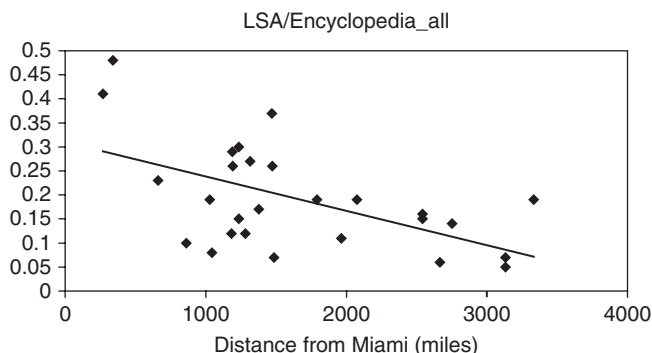


Fig. 9.1a

In addition to L–R associations, the brain forms L–L and R–R associations. Again, Hebbian learning is the mechanism. Computational techniques of large corpora of text, such as *latent semantic analysis* (LSA; Landauer and Dumais 1997), can be used to uncover patterns of L–L co-occurrences. These co-occurrences can be temporal or spatial (linguistic constructions occurring in similar contexts, without necessarily temporarily co-occurring in such contexts). Just as linguistic traces can co-occur in time and space, so referential traces can co-occur in time and space (Zwaan and Madden 2005). For example, tomatoes and lettuce can be found together in salads; staplers and notepads on desktops; and guitars and drums in recording studios. Given that language most often describes situations in the real world, or in realistic environments, L–L co-occurrences tend to reflect R–R co-occurrences.

Techniques such as LSA can capitalize on such second-order correlations between co-occurrences. For example, Figure 9.1a shows the correlation ($r = 0.55$) between the distance from Miami for 27 cities in the United States and the LSA cosine in the

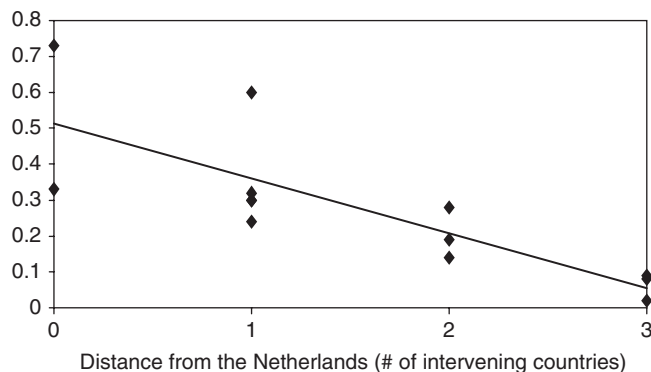


Fig. 9.1b

encyclopaedia corpus (<http://lsa.colorado.edu/>).¹ Using the same corpus, Figure 9.1b shows the correlation between the ordinal distance from the Netherlands – 0 for bordering countries, 1 for countries separated from the Netherlands by one other country, and so on – and the cosines between them ($r = 0.79$).

Similar substantial correlations can be found for states in the US and South American countries. These correlations presumably reflect the fact that cities or countries in close spatial proximity are more likely to be described in similar contexts than more remote cities or countries. This could be for a variety of reasons including geological, geographical, and historical overlap, and the greater likelihood of interactions via trade, travel, and war: all of these are common topics in encyclopaedias. Although impressive, and potentially useful, such second-order correlations do not allow the inference that meaning is grounded in co-occurrences of linguistic constructions. Rather, they suggest that L–L associations often mirror R–R associations and, due to this correspondence, can be used to enhance language comprehension, for example with regard to enhancing the ability to anticipate upcoming information. The notion of anticipation in comprehension will be developed below.

9.1.1 Cross-modal activations

Given that experiential traces are multimodal representations, it follows that when a trace is activated in one modality, its components in other modalities resonate as well. For example, phonological representations are activated during reading (e.g., Stone *et al.* 1997) and orthographic representations are activated during speech comprehension (e.g., Seidenberg and Tanenhaus 1979) and speech production (Damian and Bowers 2003). Furthermore, there is evidence that auditory representations of words activate the motor programs used to produce these words and viewing letters activates, in expert typists, the motor programs used to type them (Rieger 2004; for related evidence see Beilock, *in press*; van den Bergh *et al.* 1990). By the same token, aspects of experience in one modality may facilitate retrieval of other aspects of the experience. Classic findings in the memory literature of encoding specificity are cases in point. More recently, it has been shown that proprioceptive cues facilitate retrieval of posture-congruent autobiographical memories (Dijkstra *et al.* 2007). For instance, assuming a reclining posture facilitates retrieval of memories of visits to the dentist.

9.2 Resonance

Linguistic traces resonate with linguistic input, causing associated referential traces to resonate as well. The general idea here is consistent with an instance-based view of memory (e.g., Hintzman 1986), according to which experiential traces resonate to varying degrees with the input.

¹ See Kintsch (Chapter 8, this volume) and Louwerse and Jeuniaux (Chapter 15) for similar analyses and somewhat different conclusions.

The extent to which experiential traces resonate is a function of their similarity to the input and their current level of activation. For example, if a referential representation has recently been activated, it is likely to resonate with a later occurring word, thus inserting itself in the mental simulation. In support of this idea, we recently found that incidental exposure to word–picture combinations can affect later reading (Aveyard *et al.*, submitted). The experiment involved two phases: a word–picture verification phase presenting critical objects in specified shapes or conditions (e.g., a perched eagle), and a subsequent reading phase (ostensibly unrelated to the first) in which the subjects' eye movements were tracked. In this phase, texts were presented that implicitly disambiguated shapes or conditions of critical objects (e.g., 'In the sky an eagle ...'). Reading times on the target word ('eagle') were longer when the implied shape in the text mismatched the associated shape from the first phase (for example, a perched eagle cannot be found in the sky). Furthermore, total reading times on the prepositional phrase were longer in the mismatch condition than in the match condition, presumably indicating that the activation of the contextually inappropriate shape caused subjects to re-encode the location of the target entity in an attempt to resolve the inconsistency.

Abstractions are an automatic by-product of the resonance process. For example, the word 'not' will activate all instances of having experienced the input "not" (e.g., 'Don't pet the doggie;' 'Do not leave your clothes on the floor;' 'Do not talk in class'). Presumably, the only thing these traces have in common is that a state of the environment has triggered the activation of a goal, the execution of which is subsequently thwarted. This abstraction process (see also Zwaan and Madden 2005), which may involve a different set of traces each time, renders the meaning of 'not' as a sequence of activation patterns. First, a representation of an expected or desired state is activated, subsequently activating the motor programs needed to bring about that state. The execution of these programs is then blocked, which likely results in an emotional response (e.g., surprise, annoyance, anger, frustration). Thus, the difference between concrete and abstract words is that concrete words activate a relatively homogeneous set of traces, whereas abstract words activate a more heterogeneous set of traces, leading to patterns of activation that differ more across and within individuals than traces of concrete concepts (see also Barsalou and Wiemer-Hastings 2005).

9.3 Mental simulation

Activated experiential traces are used in mental simulations of the described situation (Barsalou 1999).² A notion of mental simulation that is relevant to language comprehension has been developed in research on motor control and in research on action observation.

Anticipation is a crucial mechanism in motor control. Given the inherent delays in perception – for example, it takes about one hundred milliseconds to process visual feedback – and in action execution, perceptual feedback is not very useful in the control of

² Elsewhere, I have shown how this view can be viewed as a logical extension of earlier work on situation models (Zwaan 2004; Zwaan and Radvansky 1998).

motor actions. Instead, the central nervous system produces a mental simulation, called a *forward model*, which uses a copy of the efferent signal to the effectors to predict the sensorimotor consequences of motor commands. Discrepancies between the predicted and the observed perceptual input are used to fine-tune the system, by inducing synaptic weight changes, such that future predictions in similar situations will be more accurate. In one model, simulation involves a range of motor programs, which are activated in resonance fashion to the extent that they match the current situation (Wolpert *et al.* 2003).

Similar anticipatory mental simulations are thought to underlie action understanding and social interaction (Gallese *et al.* 2003; Wolpert *et al.* 2003). Empirical evidence is consistent with this idea. For example, during the performance and observation of a block-stacking task, the observer's pattern of eye fixations mirrors that of the performer; both precede the movement of the blocks (Flanagan and Johansson 2003). This experiment suggests that the observer is using the same anticipatory eye movement program to understand the action that the actor is using to perform the action. Importantly, this anticipatory pattern is not found when the actor is hidden from view and the blocks seem to move by themselves. In this case, eye movements trail the blocks. In other words, this experiment shows very elegantly that anticipatory processing occurs during action observation, but only when the observer interprets the actions as being performed by a conspecific.

Single-cell recordings in the macaque monkey ventral premotor cortex provide detailed analyses of the recruitment of motor programs during action observation (e.g., DiPellegrino *et al.* 1992). In this model, some fire when the monkey observes an action being performed that it also has in its own action repertoire (e.g., grasping a food item). These neurons have been termed *mirror neurons*. Mirror neurons have also been shown to fire when the monkey hears a sound associated with an action in its repertoire, for instance cracking a nut (Köhler *et al.* 2002).

Importantly, mirror neurons are responsive to an understanding of the goal of an action. When the monkey knew there was food behind a screen, its mirror neurons responded when the experimenter's hand moved towards the food, even though the hand disappeared behind a screen. The activation pattern was similar to a condition without the screen; some mirror neurons responded equally strongly in both conditions, whereas others responded more strongly in the full vision condition. In contrast, mirror neuron activity did not occur, with or without screen, if there was no food, but the experimenter made the same grasping movement (Umiltà *et al.* 2001). Thus, in the monkey brain, having a mental representation of the goal of a grasping action seems both necessary and sufficient for mirror neuron activation; however, this restriction does not seem to hold for the human brain (Grèzes, Armony *et al.* 2003).

A recent computational approach (Keysers and Perrett 2004) suggests how sensory information becomes associated with motor programs due to the anatomical connections between the superior temporal sulcus area, which responds to visual and auditory stimulation, and areas PF and F5, which receive input from superior temporal sulcus. Because a subset of neurons in these areas show some degree of viewpoint independence, the monkey learns to associate not only the sights and sounds of its own actions with

motor programs, but also the sights and sounds of the same actions performed by others. Converging evidence has been provided in brain imaging studies of humans (corresponding human areas are Brodmann areas 44 and 6, posterior parietal lobe, and superior temporal sulcus). When humans observe a facial action that is within their repertoire (e.g., human or monkey lip smacking), blood flow increases in the premotor cortex. However, when a facial action is observed that is outside the human repertoire (e.g., barking), activation of the visual, but not premotor, cortex occurs (Buccino *et al.* 2004). The human mirror system appears to be more flexible than the monkey's in that it responds to a broader range of actions, including mimed ones in which no goal object is present, as well as to the visual presentation of manipulable objects (Grèzes, Armony *et al.* 2003).

If motor resonance occurs during the direct observation of actions, then one might also hypothesize that it is used in the understanding of actions that are conveyed via language (e.g., Gallese and Lakoff 2005; Rizzolatti and Arbib 1998). This hypothesis predicts that exposure to words denoting or sentences describing actions or objects associated with actions would result in motor resonance. Neuroimaging (for a review, see Pulvermüller, Chapter 6, this volume) as well as behavioural research (e.g., Bub *et al.*, in press; Glenberg and Kaschak 2002; Tucker and Ellis 2003; Zwaan and Taylor 2006) supports this prediction (see Fischer and Zwaan, in press, for an extensive review of this literature).

9.4 Presonance

What is all this activation good for? It allows organisms to anticipate to a certain degree upcoming states of the environment and of the organism itself. Elsewhere, we have called this anticipatory skill, *presonance* (Zwaan and Kaschak, in press). The notion of presonance captures two components of anticipation: prediction and resonance. Experiential traces are activated by way of resonance with a stimulus in a manner that allows the organism to maximize its ability to anticipate future states of the environment and of itself. Presonance should not be thought of as slow and deliberate forecasting. Rather, it should be thought of as *fluency* – it is fast and involuntary. The idea of anticipation as a highly adaptive cognitive mechanism has been proposed by various other researchers (e.g., Berthoz 2000; Knoblich and Wilson 2005; Wolpert *et al.* 2003; Zacks *et al.*, in press).

In language, presonance takes on a dual form. Anticipatory processes operate at two different levels. First, they operate at the linguistic level. For example, co-articulation is an anticipatory process at work in language production. The way in which a phoneme is articulated is influenced by the gesture our articulatory system makes in anticipation of the phonemes that follow. There is a long history in psycholinguistics on research into syntactic strategies which allow the language processing system to anticipate upcoming constituents, thereby minimizing processing effort, except in relatively rare cases of syntactic ambiguity (e.g., Frazier 1979). Similar resource-efficient anticipatory strategies have been identified at the level of connected discourse (e.g., Kintsch and van Dijk 1978; van Dijk and Kintsch 1983). The language processing system uses these strategies to anticipate referents in upcoming clauses, thus facilitating the processing of upcoming clauses.

The idea that language processing involves the rapid use of information to anticipate likely next states in the linguistic input is consistent with constraint-satisfaction views of language processing (e.g., MacDonald *et al.* 1994; McRae *et al.* 1998; Jurafsky 1996). According to constraint satisfaction models, sentence processing proceeds by activating many possible interpretations for the sentence. These interpretations compete for activation on the basis of probabilistic information from the comprehender's experience with language. The likelihood of a particular word being used in a particular syntactic function (MacDonald *et al.* 1994), the likelihood of a particular syntactic structure being used (Jurafsky 1996), the preceding context (e.g., van Berkum *et al.* 2005; Spivey and Tanenhaus 1998), and other such factors simultaneously constrain the unfolding interpretation of the sentence.

The second level at which anticipatory processes operate is that of the referential situation. Comprehenders generate anticipations on their experience with events in the real world, for example, with regard to the chronology and temporal contiguity of the described events (e.g., Zwaan 1996), with regard to the goals and plans of story protagonists (e.g., Graesser *et al.* 1994), and with regard to a number of other situational dimensions (Zwaan *et al.* 1995; Zwaan and Radvansky 1998). Shifts on these dimensions lead to increases in on-line processing load (Zwaan *et al.* 1995; Zwaan *et al.* 1998), presumably because they bring about a decrease in predictability.

Two questions can be asked regarding the relation between presonance and language comprehension: (1) what can presonance do for language comprehension, and (2) what can language comprehension do for presonance? A plausible answer to the first question is: to facilitate the processing of incoming information, thereby enhancing the fluency of the comprehension (or production) process and minimizing the demand on cognitive resources. A plausible answer to the question of what language comprehension can do for presonance might be: to acquire, without any physical danger to the organism, anticipatory mechanisms that can be used in real-world situations. Suppose someone tells you 'Don't eat those red berries or you'll get very sick.' This knowledge allows you to anticipate what will happen when you eat the berries, and will in all likelihood prevent you from trying to eat them in the first place. Importantly, language comprehension can result in anticipatory processes that operate at a much finer level of granularity. This is a promising way of viewing the role of language comprehension because it provides an answer to the question: what is comprehension for?³

9.5 Is comprehension ornamental or instrumental?

Even if one is convinced by the rapidly growing amount of empirical evidence that perceptual and motor representations are activated during language use, several significant obstacles to embodied theories of language use remain. I will address two of them.

Even if perceptual and motor activation occur routinely during language comprehension, it does not immediately follow that they are *necessary* for comprehension.

³ Glenberg (1997) raised – and proposed an answer to – the same question about memory.

They could be ornamental, or epiphenomenal, rather than instrumental to comprehension. At the theoretical level, the epiphenomenon question can be tackled by making an appeal to parsimony. If a theory based on perceptual and motor representations can account for all the findings that amodal, abstract, and arbitrary (AAA) symbol systems can account for, then why should we postulate the latter, given the wealth of evidence for the former and the paucity of evidence for the latter (Barsalou 1999)? Moreover, if perceptual and motor representations are routinely activated during comprehension, but do not play a meaningful role, we can ask, echoing Wilson and Knoblich (2005): ‘What is the purpose of this neurological extravaganza?’ It does not seem theoretically parsimonious to assume that there are AAA symbols that are doing all the work and that the embodied representations are just along for the ride. Moreover, no brain region has been found that features as the locus for converting analogue perceptual information into abstract representations (Barsalou 1999).

However, none of this should be taken to imply that lexical representations and the associations between them do not have a role to play in an embodied view of language use. Lexical representations are just another type of experiential trace. Moreover, lexical (and sublexical) associations augment comprehension in at least three ways. First, they enhance our ability to anticipate upcoming information – to pre-sonate – by providing an additional stream of anticipatory information. Not only do we anticipate upcoming information by activating relevant situational information, we also generate anticipations based on lexical associations and typical constituent sequences. Second, lexical associations allow us to create new experiential traces by way of referential bootstrapping. For example, linguistic descriptions help us form experiential traces of unfamiliar animals, such as okapis, or nonexistent animals, such as griffins, by instructing us how to use L–L combinations to create R–R combinations via the already stored L–R associations. Third, L–L associations lower the demands on attention and short-term memory. For example, if we have to memorize a sequence of actions to be performed, it may be efficient to use the associated words as a shorthand.

9.6 What’s next?

Several questions should be high on the research agenda for the coming years. Among them are the following:

- (1) in addition to showing that perceptual and motor representations are activated in language comprehension, we need to specify when and how this occurs,
- (2) we need to examine whether and how perceptual and motor simulations interact in language comprehension,
- (3) we need to develop theories about different levels of comprehension. For example, deeper comprehension should be attainable if we have the described actions in our repertoire than if we do not, and
- (4) we need to examine more closely the role of prior experience (not just declarative knowledge) in language comprehension.

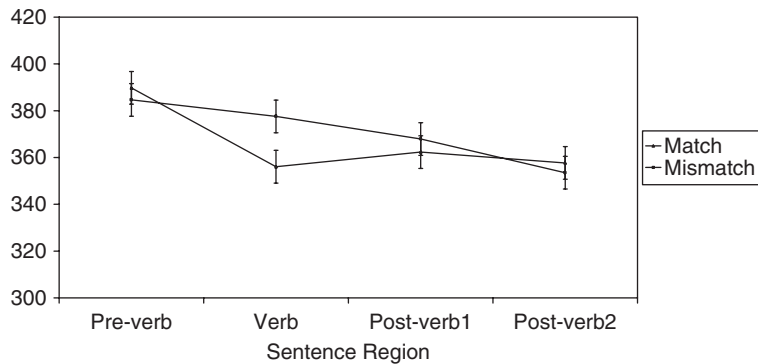


Fig. 9.2

In my laboratory, we have begun addressing some of these questions. I will conclude with a brief description of a series of experiments aimed at examining the waxing and waning of motor resonance during sentence comprehension. Glenberg and Kaschak (2002) showed that motor resonance occurs when subjects make sensibility judgments about sentences ('does this sentence make sense, yes or no?'). However, this does not directly implicate motor resonance in comprehension. In order to address this question, we asked subjects to read sentences one segment (two or three words) at a time by rotating a knob, with every four degrees of rotation resulting in the presentation of a new segment (Zwaan and Taylor, 2006; experiment 4).⁴

The critical sentences implied a manual rotation action, for example, 'After/ lighting/ the candles/ for the/ romantic/ evening,/ he/ dimmed/ the/ lights.' Each sentence had a critical region ('dimmed' in the example). We hypothesized that reading times on the critical region should be longer if the actual and implied rotation were in different directions – for example, turning down the volume requires counterclockwise motion and so should be read more slowly when read via clockwise rotation than when read via counterclockwise rotation. The idea that motor resonance should occur immediately is consistent with constraint-based theories of sentence comprehension and with the view outlined here. Also of interest was the question of what would happen to motor resonance in the regions after the critical region – an article and a noun in all critical sentences. Figure 9.2 shows the reading times per segment for the critical sentences, the regions before the verb having been collapsed into one.

The results show motor resonance first appearing on the critical verb and then disappearing after the verb. The immediate occurrence of motor resonance relative to the semantic content of the sentence was in accordance with our expectations. However, the

⁴ Reading by turning a knob at first sight may seem awkward. However, as pilot testing revealed, in fact it feels surprisingly natural, especially compared with advancing through the sentence via key presses.

disappearance of the effect after the verb is intriguing. One potential explanation of this finding is that motor resonance only occurs when the action is what is focused on by the sentence: we call this the *linguistic focus hypothesis* (LFH). In a recent series of experiments (Taylor and Zwaan, in press), we put the LFH to the test. The sentences in the previous experiment all shifted attention away from the performance of the action itself to the acted-upon object (e.g., turning back the clock, dimming the light, screwing in a lightbulb, turning down the volume).⁵ According to the LFH, if a sentence maintains focus on the action itself, for example, by modifying it with an adverb, motor resonance should extend beyond the verb. We tested this prediction by using sentence pairs of the following type: ‘The runner/ was very/ thirsty./ A fan/ handed him/ a bottle/ of cold/ water/ which he/ opened/ quickly.’ According to the LFH, motor resonance should now occur on both the verb and the adverb, and the results support this prediction (Taylor and Zwaan, in press; experiment 1).

We next performed an even stronger test of the LFH, replacing the action-modifying adverbs from the previous experiment with subject-modifying adverbs. Consider for example: ‘The runner/ was very/ thirsty./ A fan/ handed him/ a bottle/ of cold/ water/ which he/ opened/ eagerly.’ The adverb in this sentence tells us more about the runner’s mental and physical state than about the manner in which the action is performed. Therefore, the LFH predicts that motor resonance should occur on the verb, but not on the adverb. The results support this prediction (Taylor and Zwaan, in press; experiment 2).

We believe that these, and eventually much more fine-grained analyses, are needed to achieve a better insight into how language orchestrates mental simulations in the comprehender.

9.7 Discussion

To summarize, I have tried to outline a view of language comprehension that is decidedly embodied. However, I have also tried to show that this view is not as radical a departure from earlier disembodied views as might be believed at first sight. There still is an important role for lexical representations and their associations. However, these are now conceptualized as multimodal embodied mental representations, or experiential traces. I have identified pre-resonance as a crucial mechanism in motor control, action observation, and language comprehension. Finally, I have outlined several topics that should be high on the research agenda, and have briefly sketched how we are currently attempting to address one of them. These activities lead me to conclude that instead of asking *whether* cognition, and therefore language comprehension, is embodied, it is much more fruitful to ask *how* it is embodied.

Debate

The following is a simulated discussion of this chapter. Because of an unfortunate snafu, Rolf Zwaan was unable to attend the meeting in Garachico. The editors asked

⁵ See MacWhinney (2005) on grammatical constituents and perspective shifts.

Max Louwerse to prepare questions of the sort he would have asked if the paper had been presented at Garachico, and we asked Rolf Zwaan to respond to those questions.

Max Louwerse: Overview papers like these remind me yet again how innovative the embodied cognition literature is. I am glad that you agree that the question *how* language comprehension is embodied is more fruitful than the question *whether* language comprehension is embodied. I think, in part, thanks to the embodiment paradigm reigning over the last decade, the less fruitful question has already been answered: language comprehension can be embodied. We may, however, disagree on the modal auxiliaries: *can* language comprehension be embodied or *must* it (always) be? You make the argument that (L–R) associations are formed between linguistic traces and referential traces, which establish the link between symbolic processing and embodied processing. You also argue for R–R associations and L–L associations. Computational linguistic techniques like LSA illustrate these L–L associations. These L–L associations are important but merely a by-product of R–R associations. And that’s where our points of view start to differ. I would therefore like to raise four questions.

First, if Hebbian learning explains L–L, R–R, and L–R associations, why is it the case that for every linguistic trace a referential trace must be activated in order to assign meaning? We already agreed that, in principle, symbol grounding is needed. What is your view on a hypothesis that would state that some L–R associations are formed, which then allow meaning to be bootstrapped through L–L associations? In other words, could it be the case that some symbol grounding may be needed under certain circumstances, but not all symbol grounding under all circumstances?

Rolf Zwaan: One important question that needs to be addressed first is whether R representations are activated routinely during language processing (production and comprehension, but in my chapter I focus on comprehension). My answer to this question would be ‘yes.’ Evidence is coming out all of the time that supports this view. I believe you’re asking whether this activation is necessary for comprehension. One way to examine this is to see what happens if the relevant brain systems cannot be used for comprehension, for example, because they are otherwise engaged (e.g., in a dual-task paradigm) or because they are incapacitated (either permanently as a result of lesions or temporarily as a result of the administration of transcranial magnetic stimulation). Absolutely key here is to use or develop comprehension tests that are sufficiently sensitive. Researchers are beginning to address this issue, but it is still too early in the game to confidently draw conclusions, although there certainly is some compelling evidence already that provides an affirmative answer to the necessity question. Your hypothesis ‘that some symbol grounding may be needed under certain circumstances, but not all symbol grounding under all circumstances’ doesn’t seem very testable to me. In order to begin answering it, you would have to specify the conditions under which you expect symbol grounding to be necessary and those under which you do not expect symbol grounding to be necessary.

Louwerse: Second, you argue that because perceptual and motor representations are routinely activated during comprehension, they must play a meaningful role

(‘what would otherwise be the purpose of this neurological extravaganza?’). I would like to pose the same question to you: if language structures show to encode perceptual and motor representations and we routinely use language, they must play a meaningful role in cognition – what would otherwise be the purpose of this linguistic extravaganza? Are these linguistic structures simply a by-product of embodied cognition, or are they in fact used by comprehenders as a mnemonic shortcut to meaning?

Zwaan: I’m not sure I understand what you mean when you say that ‘language structures ... encode perceptual and motor representations.’ The way I conceptualize it, language structures activate perceptual and motor representations. They therefore play a very important role in comprehension, because they orchestrate the activation and integration of these representations.

Louwerse: Third, you allude to the claim that the ‘theory based on perceptual and motor representations can account for all the findings that AAA symbol systems can account for.’ We could discuss whether amodal and abstract symbol systems like the language system are truly arbitrary, but I have a more mundane question. I have not found evidence that embodiment theories can account for all the findings that AAA symbol systems can account for. I will not give an overview of the evidence for AAA symbol systems (see Landauer *et al.* 2007), but the two studies I know that attempt to account for abstract words are the ones you refer to: Zwaan and Madden (2005) and Barsalou and Wiemer-Hastings (2005). To call these exploratory studies (I use Barsalou and Wiemer-Hastings’ description) evidence that embodiment theories can account for all the findings that AAA symbol systems can account for is a stretch. Moreover, I am tempted to make the opposite argument: in various cases my lab has found evidence that language structures can explain embodiment findings (see Louwerse 2007; Louwerse *et al.* 2006, Louwerse and Jeuniaux, Chapter 15, this volume). So what would be your view on a position stating that a theory based on perceptual and motor representations can account for some of the findings that AAA symbol systems cannot account for, and that a theory based on amodal and abstract symbolic representations can account for other findings that perceptual and motor representations cannot account for?

Zwaan: The easy answer would be that amodal models account for concrete and abstract representations equally badly. They merely transcode the input. For sure, embodied frameworks are currently doing a much better job accounting for concrete concepts than they do for abstract concepts. But they have made promising ways to address this shortcoming, Barsalou’s approach being one of them. But you are obviously right that a lot more work is needed. I strongly disagree with your self-assessment that the research you report in the book chapters you mention can explain embodiment findings that have been published in the literature. What your LSA analyses do, and cleverly so, is mimic the output of selected embodiment experiments. This does not imply that they mimic the underlying cognitive processes. To me, this is the fatal flaw in your approach. Your reasoning amounts to the claim that Deep Blue provides a model of a human chess grandmaster.

Louwerse: Finally, I would like to conclude with a broader, overarching question. Psycholinguistic research was dominated by symbolic theories in the 1960s to the 1990s. It now seems to be dominated by embodiment theories, from the 1990s to 2007. Are we perhaps just dealing with mood swings from our colleagues, who move explanations from one (symbolic) side of the spectrum to the other (embodied) side, while in all truth the answer lies – blatantly obvious – right in the middle?

Zwaan: I'm not sure I agree that psycholinguistic research is 'dominated by embodiment theories.' The situation is quite the opposite, actually. In fact I'd go so far as to say that even the name 'psycholinguistics,' with its focus on syntax, is inherently nonembodied. If we take the area of cognition more generally, however, it is clear that there is an increasing interest in embodiment. In large part this is because of the influence of neuroscience, which keeps on piling up evidence about cognitive processes that seem difficult to explain from a traditional framework, but that can much more easily be explained by embodied theories. However, I'd still hesitate to say that the cognitive literature is dominated by embodiment theories.

I don't think we're dealing with mood swings. Rather, I think we're dealing with a larger-scale change in the field. Symbolic theories from the 1960s to the 1990s were inspired by the computer metaphor of cognition and by developments in computer science and linguistics. The brain was largely irrelevant to these theories. The advent of neuroimaging techniques, which provide a window on 'the brain in action' as it performs cognitive tasks, has made it increasingly clear that theories of cognition cannot simply ignore the brain and that the brain is not neatly organized in different modules, each with their own jobs to do, as traditional cognitive theories would have it. So I see the change to embodied approaches as a biologically rather than linguistic-philosophically inspired one. I suspect that this trend will continue.

Author note

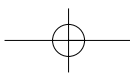
This research is supported by National Institutes for Health grant MH-63972 and National Science Foundation grant BCS-0446637. I thank Tobias Richter, Larry Taylor, and Rich Yaxley for helpful comments on a previous version. Please address all correspondence regarding this paper to: Rolf A Zwaan, Department of Psychology, Florida State University, Tallahassee, FL 32306-1270, USA. Email may be sent to: zwaan@psy.fsu.edu.

References

- Barsalou LW (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–660.
- Barsalou LW, Wiemer-Hastings K (2005). Situating abstract concepts. In D Pecher, R Zwaan, Eds. *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thought* (pp. 129–63). New York, NY: Cambridge University Press.
- Beilock SL, Holt LE (in press). Embodied preference judgments: can likeability be driven by the motor system? *Psychological Science*.
- Berthoz A (2000). *The Brain's Sense of Movement*. Cambridge, MA: Harvard University Press.
- Bub DN, Masson MEJ, Cree GS (in press). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition*.

- Buccino G, Binkofski F, Fink GR *et al.* (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, 13, 400–4.
- Damasio AR (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York, NY: Harcourt Brace.
- Damian MF, Bowers JS (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language*, 49, 119–32.
- DiPellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, 91, 176–80.
- Dijkstra K, Kaschak MP, Zwaan RA (2007). Body posture affects retrieval from autobiographical memory. *Cognition*, 102, 139–49.
- Fischer MH, Zwaan RA (in press). Embodied language: a review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology*.
- Flanagan JR, Johansson RS (2003). Action plans used in action observation. *Nature*, 424, 769–71.
- Frazier L (1979). *On Comprehending Sentences: Syntactic Parsing Strategies* [doctoral thesis]. Storrs, CT: University of Connecticut.
- Gallese V, Lakoff G (2005) The brain's concepts: the role of the sensory-motor system in reason and language. *Cognitive Neuropsychology*, 22, 455–79
- Glenberg AM (1997). What memory is for. *Behavioral and Brain Sciences*, 20, 1–55.
- Glenberg AM, Kaschak MP (2002). Grounding language in action. *Psychonomic Bulletin and Review*, 9, 558–65.
- Graesser AC, Singer M, Trabasso T (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101, 371–95.
- Grezes J, Armony JL, Rowe J, Passingham RE (2003) Activations related to 'mirror' and 'canonical' neurones in the human brain. *NeuroImage*, 18, 928–37.
- Grezes J, Tucker M, Armony J, Ellis R, Passingham RE (2003). Objects automatically potentiate action: an fMRI study of implicit processing. *European Journal of Neuroscience*, 17, 2735–74.
- Hintzman DL (1986). Schema-abstraction in a multiple trace model. *Psychological Review*, 93, 411–28.
- Jurafsky D (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20, 137–94.
- Keyers C, Perrett DI (2004). Demystifying social cognition: a Hebbian perspective. *Trends in Cognitive Sciences*, 8, 501–7.
- Kintsch W, Van Dijk TA (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363–94.
- Köhler E, Keyers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297, 846–8.
- Landauer TK, Dumais ST (1997). A solution to Plato's problem: the latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211–40.
- MacDonald MC, Pearlmutter NJ, Seidenberg MS (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 89, 483–506.
- MacDonald MC (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, 9, 157–201.
- MacWhinney B (2005). The emergence of grammar from perspective taking. In D Pecher, RA Zwaan, Eds. *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking* (pp. 198–223). Cambridge: Cambridge University Press.
- McRae K, Spivey-Knowlton MJ, Tanenhaus MK (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38, 283–312.

- Rieger M (2004). Automatic keypress activation in skilled typists. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 555–65.
- Rizzolatti G, Arbib MA (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188–94.
- Sadoski M, Paivio A (2001). *Imagery and Text: A Dual Coding Theory of Reading and Writing*. Mahwah, NJ: Erlbaum.
- Seidenberg MS, Tanenhaus MK (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 546–54.
- Spivey MJ, Tanenhaus MK (1998). Syntactic ambiguity resolution in discourse: modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1521–43.
- Spivey MJ, Richardson DE (in press). Language embedded in the environment. In P Robbins, M Aydede, Eds. *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press.
- Stone GO, Vanhoy M, Van Orden GC (1997). Perception is a two-way street: feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language*, 36, 337–59.
- Taylor LJ, Zwaan RA (in press). Linguistic focus and motor resonance. *Quarterly Journal of Experimental Psychology*.
- Umiltà MA, Köhler E, Gallese V *et al.* (2001). ‘I know what you are doing’: a neurophysiological study. *Neuron*, 32, 91–101.
- Van Berkum JJA, Brown CM, Zwitserlood P, Kooijman V, Hagoort P (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 443–67.
- Van den Bergh O, Vrana S, Eelen P (1990). Letters from the heart: affective categorization of letter combinations in typists and nontypists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 16, 1153–61.
- Van Dijk TA, Kintsch W (1983). *Strategies of Discourse Comprehension*. New York, NY: Academic Press.
- Wilson M, Knoblich G (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131, 460–73.
- Zwaan RA (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 920–33.
- Zwaan RA (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1196–207.
- Zwaan RA (2004). The immersed experiencer: toward an embodied theory of language comprehension. In BH Ross, Ed. *The Psychology of Learning and Motivation*, Vol. 44 (pp. 35–62). New York, NY: Academic Press.
- Zwaan RA, Kaschak MP (in press). Language comprehension as a means of ‘re-situating’ oneself. In P Robbins, M Aydede, Eds. *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press.
- Zwaan RA, Langston MC, Graesser AC (1995). The construction of situation models in narrative comprehension: an event-indexing model. *Psychological Science*, 6, 292–7.
- Zwaan RA, Madden CJ (2005). Embodied sentence comprehension. In D Pecher, RA Zwaan, Eds. *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking* (pp. 224–45). New York, NY: Cambridge University Press.
- Zwaan RA, Magliano JP, Graesser AC (1995) Dimensions of situation-model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 386–97.
- Zwaan RA, Radvansky GA (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162–85.



Zwaan RA, Radvansky GA, Hilliard AE, Curiel JM (1998). Constructing multidimensional situation models during reading. *Scientific Studies of Reading*, 2, 199–220.

Zwaan RA, Taylor LJ (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General*, 135, 1–11.

