THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2009, 62 (6), 1108–1114

Psychology Press
Taylor & Francis Group

Short article

Language comprehenders retain implied shape and orientation of objects

Diane Pecher, Saskia van Dantzig, Rolf A. Zwaan, and René Zeelenberg
Erasmus University, Rotterdam, The Netherlands

According to theories of embodied cognition, language comprehenders simulate sensorimotor experiences to represent the meaning of what they read. Previous studies have shown that picture recognition is better if the object in the picture matches the orientation or shape implied by a preceding sentence. In order to test whether strategic imagery may explain previous findings, language comprehenders first read a list of sentences in which objects were mentioned. Only once the complete list had been read was recognition memory tested with pictures. Recognition performance was better if the orientation or shape of the object matched that implied by the sentence, both immediately after reading the complete list of sentences and after a 45-min delay. These results suggest that previously found match effects were not due to strategic imagery and show that details of sensorimotor simulations are retained over longer periods.

Keywords: Sentence processing; Sensorimotor processing; Memory.

When people comprehend language, they mentally represent the situation described by the text (Johnson-Laird, 1983; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). Several theories have proposed that sensorimotor processes are involved in representing concepts and events (Barsalou, 1999; Glenberg, 1997; Goldstone & Barsalou, 1998; Pulvermüller, 1999; Wilson, 2002; Zwaan, 2004). According to these theories, mental representations have many similarities to actual experiences and use the same or highly similar processes. There is now quite a body of data lending support to this embodied view of concept representation (e.g., Borghi, Glenberg, & Kaschak, 2004; Pecher, Zeelenberg, & Barsalou, 2003, 2004; Zwaan & Madden, 2005; Zwaan, Stanfield, & Yaxley, 2002).

In his perceptual symbols theory, Barsalou (1999) provides a detailed account of how cognition may be grounded in sensorimotor processes.

Correspondence should be addressed to Diane Pecher, Psychology Department, T12–33, Erasmus University Rotterdam, Postbus 1738, 3000 DR Rotterdam, The Netherlands. E-mail: pecher@fsw.eur.nl

This research was supported by a grant from the Netherlands Organization for Scientific Research (NWO) to Diane Pecher, a grant from the Vereniging Trustfonds EUR to Saskia van Dantzig, and grant BCS-0446637 from the National Science Foundation (NSF) to Rolf Zwaan. Part of this research was done when Rolf Zwaan was at Florida State University, Tallahassee, Florida, USA. We thank Mante Nieuwland for use of his materials and Kim Melchior for her assistance with testing subjects.

http://www.psypress.com/qjep
DOI:10.1080/17470210802633255

He proposes that the modality-specific systems that are used for perception and action are also used to represent concepts. During perception and action, the patterns of activation that result from experience are captured. These form the basis for the perceptual symbols that are used to re-enact experiences. During thought a concept is represented by perceptual symbols through a simulation of sensorimotor interaction with the concept. These simulations are partial and sketchy but nevertheless grounded in embodied experiences.

Evidence for the role of visual simulations in representing the meaning of language was obtained by Stanfield and Zwaan (2001) and Zwaan et al. (2002; see also Holt & Beilock, 2006). Stanfield and Zwaan presented sentences in which the horizontal or vertical orientation of an object was implied. For example, in the sentence *John put the pencil in the cup* it is implied that the pencil is oriented vertically. The sentence *John put the pencil in the drawer* implies a horizontal orientation. Immediately following the sentence, a picture was presented, and the participant decided whether the depicted object had been mentioned in the preceding sentence. On the trials of interest, the picture showed the object from the sentence in either the implied orientation or in the perpendicular orientation. Participants were faster and more accurate to recognize the object when the orientation of the picture matched the orientation implied by the sentence than when it did not match. Similar results were obtained by Zwaan et al. for pictures that matched or did not match the implied shape of an object (e.g., *The ranger saw the eagle in the sky*, followed by a picture of an eagle with folded wings or with outstretched wings). These results provide evidence for the idea that language comprehenders represent the meaning of a sentence by simulating a visual experience of what is described by the sentence.

In the present study we investigated whether these visual simulations are formed automatically when language comprehenders read sentences. An alternative explanation for the results of Stanfield and Zwaan (2001) and Zwaan et al. (2002) is that the specific tasks used in these studies may have motivated participants to use visual imagery. Participants may have used conscious imagery while reading the sentence in order to quickly recognize the picture as old or new. Because participants had to compare the object mentioned in the sentence to a picture that was presented immediately after the sentence, this may have motivated the strategic use of visual imagery during sentence reading. Thus, their results leave open the possibility that visual simulations are only used when the task demands strategic imagery, and that they are not automatically used during normal language comprehension.

In order to investigate this alternative explanation, we used a delayed picture recognition task. During the implicit study phase, participants were presented with sentences implying a particular orientation or shape (e.g., *The handyman made a hole in the wall with his drill*, which implies a drill in horizontal orientation). They judged the sensibleness of these sentences and a set of distractor sentences (e.g., *Wendy didn't notice that a child had crawled into her soda can*). Once the complete list of sentences had been read, pictures were presented in a surprise recognition memory task. This design ensured that, while reading the sentences, participants did not know that they would have to recognize pictures in the subsequent memory test. If participants automatically use visual simulations to comprehend the sentences, we expect their picture recognition performance to be better in the match than in the mismatch condition. Memory performance should be affected by the overlap between the mental representation at study and the physical appearance of the stimuli presented at test, as is predicted by the transfer-appropriate processing principle (Durgunoğlu & Roediger, 1987; Morris, Bransford, & Franks, 1977). If the previous results were due to strategic visual imagery that was induced by the task rather than by normal reading processes, however, there should be no difference in performance between the match and mismatch conditions. In order to investigate whether such a match effect remains after a longer delay, we varied the delay between sentence reading task and recognition test. One group of participants was tested immediately after the sentence reading task, and another group of participants was tested after an unrelated intervening task that lasted about 45 min.

Method

Participants

A total of 92 students at the Erasmus University Rotterdam participated for a small monetary reward. The participants were recruited for a larger package of experiments that lasted about 75 min. Of the participants, 48 received the recognition test immediately after the sentence judgement task; the remaining 44 received the recognition task after a 45-min delay.

Materials

A set of 92 sentence–picture quadruplets was created similar to the materials used in Stanfield and Zwaan (2001) and Zwaan, Stanfield, and Yaxley (2002). Each quadruplet consisted of two sentences and two black-and-white pictures, such that one of the pictures matched the implied shape or orientation of the object in one of the sentences, and the other picture matched the implied shape or orientation of the object in the other sentence. For example, a picture of a closed tube of toothpaste was the matching picture for the sentence *Angela put the toothpaste in her shopping basket*, and a picture of an opened tube of toothpaste was the matching picture for the sentence *Angela put the toothpaste on her toothbrush*. In 40 quadruplets the implied shape was manipulated, and in 52 quadruplets the implied orientation was manipulated.

Each participant received only one sentence and one picture from each quadruplet. Four counterbalanced versions were created, so that each possible combination of sentence and picture from a quadruplet was used. Note that each sentence and each picture was used both in the match condition and in the mismatch condition across the four versions. Each counterbalanced version had an equal number of match and mismatch pairs. A filler set of 92 nonsense sentences (e.g., *The hiker had a pocket knife with all kinds of veterans*) was created for the sentence judgement task during the study phase. In addition, a filler set of 92 pictures was created. These pictures served as nonstudied distractors on the recognition memory test. All sentences were Dutch. The pictures were selected from various sources (Bonin, Peereman, Malardier, Meot, & Chalard, 2003; Stanfield & Zwaan, 2001; Starreveld, 2000; Zwaan et al., 2002) or created from images found on the internet. For the orientation condition the pictures were rotated in order to present objects at two different orientations. For the shape condition different pictures were used that represented the same object in different shapes (e.g., an opened and a closed tube of toothpaste).

Procedure

In the study phase participants performed a sensibility judgement task on the sentences. In the test phase they performed a surprise recognition memory task on the pictures. For 44 participants, an unrelated experiment that lasted about 45 min intervened between the study phase and the test phase. For the other 48 participants the test phase immediately followed the study phase. To control for possible effects of position in the experiment package, 24 participants performed both tasks before the unrelated experiment, and 24 participants performed both tasks after the unrelated experiment.

All 92 experimental and 92 filler sentences in the study phase were presented in random order. Each trial started with the presentation of a fixation point (***) for 500 ms. Then the sentence was presented in the centre of the screen until a response was given or for a maximum of 4,000 ms. The participant responded by pressing the z-key with their left index finger for a *no*-response or the m-key with their right index finger for a *yes*-response. If the response was incorrect or too late, feedback—*Fout* (*Error*) or *Te laat* (*Too late*), respectively—was given for 1,000 ms. If the response was correct, the following trial started immediately. After 92 trials there was a short break in which feedback was given on the participant's performance in the first block. If the error percentage exceeded 15%, the participant was urged to respond more accurately. If the error percentage was lower than 5%, the participant was complimented.

In the short delay condition the recognition memory task was given after a short break (in which the participant notified the experimenter that the sentence task had finished, and the experimenter started the recognition task) that lasted less than 2 min. In the long delay condition the recognition task was given after the participant had participated in an unrelated experiment that lasted between 40 and 50 min. Participants were told that they would see a sequence of pictures, and that their task was to judge, for each, whether the object in the picture had been in one of the sentences that they had read in the study phase. Each trial started with the presentation of a fixation point (***) for 500 ms on the computer screen. Then the picture was presented in the centre of the screen and remained visible until the participant responded. Participants decided whether the depicted object had been presented in one of the sentences by pressing the m ("old") or z ("new") key on the computer keyboard. The pictures were presented in random order. At the end of the task feedback on the participant's accuracy was given.

The unrelated intervening task consisted of an in-basket task such as used for job assessments. This was part of an unrelated experiment that investigated performance in teamwork. Participants worked on this task in groups of three. They received memos and had to make team decisions about how to manage the memos—for example, make decisions about a company's inventory. After completing the group task, participants filled out two questionnaires about the team performance.

Results

Mean accuracy per condition was calculated for each participant. Data from 5 participants were replaced by those from 5 new participants, because their mean accuracy was at or below chance ($d' \leq 0$). The mean hit rates (i.e., the percentage of "old" responses to studied concepts) in the recognition memory task are presented in Table 1. The d' scores (a measure of sensitivity or memory strength: see MacMillan & Creelman, 1991) are presented in Figure 1. A 2 (shape vs. orientation) × 2 (match vs. mismatch) × 2 (short vs. long delay) repeated measures ANOVA was performed on the d' scores. As predicted by the perceptual simulation hypothesis, d's were higher for pictures that matched the implied shape or orientation than for pictures that did not match the implied shape or orientation, $F(1, 90) = 7.15$, $p < .01$. There were no interaction effects between match and the other variables, all Fs < 1. Memory strength was higher for pictures from the orientation condition than for pictures from the shape condition, $F(1, 90) = 11.18$, $p < .01$. The effect of delay approached significance, $F(1, 90) = 3.86$, $p = .052$, indicating that performance was better after the short than after the long delay.

Table 1. *Mean hit rates in the picture recognition task as a function of match, type of match, and delay between study and test*

		Delay	
Type		*Short*	*Long*
Orientation	Match	.59	.56
	Mismatch	.56	.54
Shape	Match	.56	.52
	Mismatch	.52	.48

Note: The false alarm rate (i.e., the percentage of "old" responses to nonstudied concepts) was .18 in the short delay condition and .20 in the long delay condition.

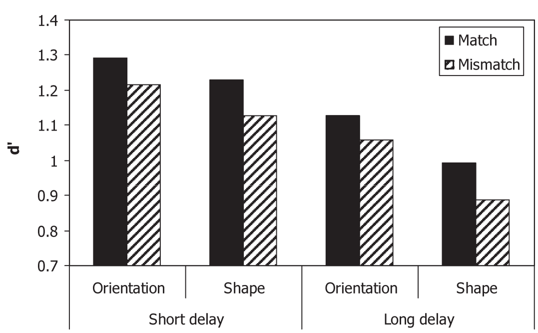

Figure 1. *Mean* d′*s in the picture recognition task as a function of match, type of match, and delay between study and test.*

Discussion

Recognition memory for pictures was better if the picture matched the implied shape or orientation of the object in an earlier sentence. These results are in line with those of Stanfield and Zwaan (2001) and Zwaan et al. (2002). In those studies, a picture followed the sentence immediately, and performance was better if the picture matched the orientation or shape of the object that was implied by the sentence. The present study shows that an effect of match can be found when sentence reading and picture recognition are separated in time.

The effect of sentence processing on immediate picture recognition has often been cited as evidence for visual simulations during online language comprehension. It is possible, however, to explain such results by strategies that are induced by the task requirements. Because in previous studies each sentence was followed directly by a picture, and the task required comparison of the picture with the meaning of the sentence, participants may have used strategic conscious imagery to enhance their performance. Stanfield and Zwaan (2001) conducted a post-hoc analysis that casts doubt on this explanation. If participants had adopted such a strategy, the match effect should be stronger in the second half of the experiment than in the first half, given that it takes a number of trials to develop a strategy. However, they obtained equal-sized match effects in the first and the second halves of the experiment. The present study provides even stronger evidence against the strategy explanation. Because the pictures were presented only after all sentences had been read, it is unlikely that the task elicited the type of conscious imagery strategy that might have played a role in previous studies. Given that our results show a similar match advantage as previous studies, we conclude that this effect is evidence of genuine comprehension processes during sentence processing.

The present results are explained by larger overlap between the simulation formed during sentence comprehension and the picture presented at test in the match condition than in the mismatch condition. It is possible, however, that the effect lies not in the simulation at study, but in the cue that is generated from the picture at test. More specifically, participants might use the picture as a cue to generate words or (partial) sentences, and subsequently match those against sentences in long-term memory. For example, a picture of a closed tube of toothpaste might generate words related to shopping, and these provide a better cue for the sentence *Angela put the toothpaste in her shopping basket* than for the sentence *Angela put the toothpaste on her toothbrush.* Such a strategy would require some time. As a result, if this strategy was responsible for the results of the present study, the match effect should be larger for slower than for faster responses. We performed a median split based on the response times in each condition and analysed d' scores again with response time as an additional factor (fastest responses vs. slowest responses). This analysis showed no interaction between response time and match ($F < 1$), suggesting that the match effect was not due to such a sentence-generation strategy.

Our present results are compatible with the view that sensorimotor experiences support "higher" cognitive processes such as language, memory, categorization, and problem solving (Barsalou, 1999; Glenberg, 1997; Goldstone & Barsalou, 1998; Pulvermüller, 1999). According to this view, modality-specific systems that support perception and action are also used to represent knowledge during offline processing. Barsalou (1999) has proposed that cognitive processes make use of simulations of perception and action in order to represent meaning. There is now a wide body of research supporting this view. In addition to studies that have obtained evidence for visual simulations during language comprehension (Connell, 2007; Richardson, Spivey, Barsalou, & McRae, 2003; Solomon & Barsalou, 2004; Spivey & Geng, 2001; Stanfield & Zwaan, 2001, Zwaan & Madden, 2005; Zwaan et al., 2002; Zwaan & Yaxley, 2003), studies have also obtained evidence for simulations in other modalities, such as audition, touch, and action (Borghi, 2005; Borghi et al., 2004; Bub, Masson, & Bukach, 2003; Glenberg & Kaschak, 2002; Marques, 2006; Pecher et al., 2003; Van Dantzig, Pecher, Zeelenberg, & Barsalou, 2008; Zwaan & Taylor, 2006).

Whereas these studies all showed that sensorimotor information was activated during online conceptual processing, the present study shows that memories that are formed during conceptual processing are still sensitive to overlap in sensorimotor information at later points in time. This finding is in line with previous studies that have shown effects of subtle variations in meaning that persisted over time (e.g., Anderson et al., 1976; Barclay, Bransford, Franks, McCarrell, & Nitsch, 1974; Barsalou, 1993; Pecher & Raaijmakers, 1999; Zeelenberg, Pecher, Shiffrin, & Raaijmakers, 2003). Whereas these studies did not focus on sensorimotor information, some recent studies obtained evidence that implicit memory for concepts is sensitive to the overlap of modality-specific information both within language (Pecher et al., 2004; Phelps, Macken, Barry, & Miles, 2006) and between language processing and picture processing (Pecher, Zanolie, & Zeelenberg, 2007). The present study shows that similar findings are obtained for sentence comprehension. Not only is memory for the content of sentences sensitive to modality-specific information, it is also sensitive to the specific properties of the visual simulation that was formed during the comprehension process. This finding provides further support for the view that sensorimotor systems are involved in language comprehension. As people are giving meaning to linguistic input, they form mental representations using perceptual symbols (Barsalou, 1999). As demonstrated by the present study, these representations are simulations that match actual perceptions, which suggests that the same systems are used for perception and conceptual representation.

The present study also shows that earlier representations can affect later processing of the same concepts. Even subtle variations such as the particular orientation or shape implied by a sentence are encoded in the representation of a concept. Thus, concepts are dynamic and change constantly with new experiences, even if those experiences are linguistic.

Original manuscript received 26 March 2007
Accepted revision received 9 June 2008
First published online 10 January 2009

REFERENCES

Anderson, R. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. V., & Trollip, S. R. (1976). Instantiation of general terms. *Journal of Verbal Learning and Verbal Behavior*, *15*, 667–679.

Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrell, N. S., & Nitsch, K. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior*, *13*, 471–481.

Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 30–89). Hove, UK: Lawrence Erlbaum.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences*, *22*, 577–660.

Bonin, P., Peereman, R., Malardier, N., Meot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments and Computers*, *35*, 158–167.

Borghi, A. M. (2005). Object concepts and action. In D. Pecher & R. A. Zwaan (Eds.), *The grounding of cognition: The role of perception and action in memory, language, and thinking*. Cambridge, UK: Cambridge University Press.

Borghi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition*, *32*, 863–873.

Bub, D. N., Masson, M. E. J., & Bukach, C. M. (2003). Gesturing and naming: The use of functional knowledge in object identification. *Psychological Science*, *14*, 467–472.

Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, *102*, 476–485.

Durgunoğlu, A. Y., & Roediger, H. L. (1987). Test differences in accessing bilingual memory. *Journal of Memory and Language*, *26*, 377–391.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*, 1–55.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*, 558–565.

Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition*, *65*, 231–262.

Holt, L. E., & Beilock, S. L. (2006). Expertise and its embodiment: Examining the impact of sensorimotor skill expertise on the representation of action-related text. *Psychonomic Bulletin & Review*, *13*, 694–701.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard UP.

MacMillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge, UK: Cambridge University Press.

Marques, J. F. (2006). Specialization and semantic organization: Evidence for multiple-semantics linked to sensory modalities. *Memory & Cognition*, *34*, 60–67.

Morris, D. C., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533.

Pecher, D., & Raaijmakers, J. G. W. (1999). Automatic priming effects for new associations in lexical decision and perceptual identification. *Quarterly Journal of Experimental Psychology*, *52A*, 593–614.

Pecher, D., Zanolie, K., & Zeelenberg, R. (2007). Verifying visual properties in sentence verification facilitates picture recognition memory. *Experimental Psychology*, *54*, 173–179.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying conceptual properties in different modalities produces switching costs. *Psychological Science*, *14*, 119–124.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2004). Sensorimotor simulations underlie conceptual representations: Modality-specific effects of prior activation. *Psychonomic Bulletin & Review*, *11*, 164–167.

Phelps, F. G., Macken, W. J., Barry, C., & Miles, C. (2006). Primacy of functional knowledge in semantic representations: The case of living and nonliving things. *Quarterly Journal of Experimental Psychology*, *59*, 1984–2009.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, *22*, 253–336.

Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, *27*, 767–780.

Solomon, K. O., & Barsalou, L. W. (2004). Perceptual simulation in property verification. *Memory & Cognition*, *32*, 244–259.

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research/Psychologische Forschung*, *65*, 235–241.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*, 153–156.

Starreveld, P. A. (2000). On the interpretation of onsets of auditory context effects in word production. *Journal of Memory and Language*, *42*, 497–525.

Van Dantzig, S., Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2008). Perceptual processing affects conceptual processing. *Cognitive Science*, *32*, 579–590.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*, 625–636.

Zeelenberg, R., Pecher, D., Shiffrin, R. M., & Raaijmakers, J. G. W. (2003). Semantic context effects and priming in word association. *Psychonomic Bulletin & Review*, *10*, 653–660.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 35–62). New York: Academic Press.

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. A. Zwaan (Eds.), *The grounding of cognition: The role of perception and action in memory, language, and thinking.* Cambridge, UK: Cambridge University Press.

Zwaan, R. A., & Radvansky, G. A. (1998) Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*, 168–171.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*, 1–11.

Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin and Review*, *10*, 954–958.