THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
0000, 00 (0), 1–6

Psychology Press
Taylor & Francis Group

Rapid communication
Readers routinely represent implied object rotation: The role of visual experience

Stephanie I. Wassenburg and Rolf A. Zwaan
Erasmus University Rotterdam, Rotterdam, The Netherlands

We conducted an eye tracking experiment to investigate whether prior visual experience affects later language processing. We assessed the effects of previously encountered pictures of objects with a vertical or horizontal orientation on the later reading of sentences that implied an object's orientation. First-pass reading times were longer when participants read about an implied orientation that did not match the orientation of the previously seen picture than when the orientation matched. This suggests that a picture encountered 20 min earlier and incidental to the reading task influenced reading. These results have implications for theories of reading comprehension and embodied cognition.

Keywords: Embodied cognition; Reading comprehension; Eye tracking; Mental simulation; Object orientation.

The traditional view in cognitive science has been that comprehension of language occurs through the integration of abstract, amodal, and arbitrary propositional representations (e.g., Fodor, 1975; Kintsch & Van Dijk, 1978; Pylyshyn, 1986). On this view, linguistic input is converted into propositional symbols; integration of these semantic representations and the knowledge residing in long-term memory constitute comprehension. A problem for this view is that propositions are abstractions that are not grounded in perceptual and motor experience; this is known as the symbol grounding problem (see Harnad, 1990).

Several alternatives to ungrounded cognition have been proposed (e.g., Barsalou, 1999; Harnad, 1990). Barsalou, for example, argues that conceptual knowledge is grounded in perceptual and motor systems. Perception and action involve certain patterns of activation in modal systems; residues of these patterns are captured and form the basis of perceptual symbols, which are typically schematic and partial because attention is selective. Thoughts, language, and grammar are cues for reenactment of experiences through activation and combination of these perceptual symbols; as such, mental representations are grounded in embodied experiences.

Consider the sentence *The carpenter pounded the nail into the wall.* Competent language users will agree that the nail's orientation is (more or less) horizontal, although this information is not stated explicitly in the text. Research has shown that comprehenders represent the implied orientation of objects. Participants presented with sentences

Correspondence should be addressed to Rolf A. Zwaan, Department of Psychology, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands. E-mail: zwaan@fsw.eur.nl

http://www.psypress.com/qjep DOI:10.1080/17470218.2010.502579

such as the one in our example responded more quickly to a picture of a horizontal nail than to one of a vertical nail shown immediately after the reading task, even though orientation was task irrelevant (Stanfield & Zwaan, 2001). It could be argued that these results were obtained because participants somehow strategically used an imagery task, given that each sentence was followed by a picture (but see Stanfield & Zwaan, 2001, for counterarguments). However, a recent experiment rules out this explanation. Participants first read sentences, some of which were similar to those of Stanfield and Zwaan, and received a surprise picture recognition test only after all sentences had been read. A similar match effect to that of Stanfield and Zwaan was obtained (Pecher, van Dantzig, Zwaan, & Zeelenberg, 2009). Thus, object orientation appears to be routinely represented during language comprehension. This is predicted by theories that assume that language comprehension involves the sensorimotor simulation of the referential situation (e.g., Barsalou, 1999; Glenberg & Kaschak, 2002; Zwaan, 2004).

Although the previous findings are useful in demonstrating an interaction between language comprehension and perceptual representations of object orientation, they demonstrate the effect going only in one direction: from language comprehension to picture recognition. From the perspective of reading comprehension, however, the reverse direction is more interesting. Does a recent visual exposure to an object in a specific orientation affect later language comprehension?

The current experiment examined the effect of a referent's orientation on online reading comprehension of sentences that implied the orientation of an object. We used the visual memory paradigm
Q1 (Zwaan, Adriaans, Pillay, & Aveyard, 2010). This paradigm consists of three phases. In the first phase, participants performed a word–picture verification task. For the experimental items, the target object was shown in a specific orientation: horizontal or vertical. We assumed that because an orientation was (a) relatively recently viewed and (b) associated with the target word, it made it highly likely that it was primed in the participant's visual long-term memory (V-LTM; Palmer, 1999). This assumption seems entirely reasonable, in light of the fact that people have been shown to exhibit extremely accurate long-term visual memory for states of objects. For example after having seen 2,500 pictures for over 5.5 hours, participants remember having seen a chest with an open door rather than the same chest with doors closed (Brady, Konkle, Alvarez, & Oliva, 2008).

The second phase of the experiment was a 15-min filler task. In the third phase, participants read sentences while their eye movements were being tracked. The sentences described objects in a certain location, which imply an orientation (e.g., *He pounded the nail into the wall*). The object's orientation either matched or mismatched that of the object shown in the picture earlier on. Importantly, the three phases of the experiment are presented as unrelated experiments, and post hoc interviews were performed to ensure they were experienced as such. Perceptual symbol theories assume that a mental simulation derives from activation and combination of perceptual symbols; it reflects physical characteristics of an object and has an analogue relationship with its referent (Stanfield & Zwaan, 2001). These theories, therefore, predict that reading comprehension will be slowed down when the implied orientation in a sentence does not match the orientation of the earlier seen picture of the target object compared to when it does match. It is expected that this difference occurs in the region where orientation is implied—namely, in the prepositional phrase that describes the location of the target object, which implies a certain orientation (vertical or horizontal). According the one-step model of Hagoort and van Berkum (2007), information from different modalities and background information are immediately activated during reading. Hence, we expected to observe the match effect to occur rather immediately in the eye tracking record—that is, during first-pass reading times.

Method

Participants

A total of 34 undergraduate students at the Erasmus University Rotterdam participated in the experiment. Data from 5 participants were unusable because of tracking errors. Data from 1 participant were removed because of exceedingly long fixation times (>2.5 times longer than those of the other participants). Of the 28 remaining participants, there were 14 females and 14 males (average age 20.3 years, range 18–24 years). They reported Dutch as their native language and confirmed that they did not have any reading problems such as dyslexia. Participants were rewarded with course credit.

Materials

In the first phase of the experiment, 80 word–picture items were presented; 20 items were critical, and 60 items were fillers. The 80 word–picture items were divided into two conditions; 40 Dutch words formed a match with the corresponding picture, and 40 Dutch words formed a mismatch. The match condition consisted of all the critical items and 20 noncritical items; 40 noncritical items belonged to the mismatch condition; therefore, 40 items required a "yes" response and 40 items a "no" response. Within the critical items, half the items were pictured with a horizontal orientation and half with a vertical orientation. Orientation was counterbalanced across participants by using two lists, which could be realized by rotating every picture 90° on its vertical axis. Each picture was scaled to occupy a square measuring approximately 350 × 350 pixels.

In the second phase a filler task—a mental rotation task—was administered.[1] This task lasted 15–20 min. In the third phase, 40 Dutch sentences were used. The critical words from the first phase were used in 20 sentences; therefore these were the critical sentences. Each sentence had the same structure: subject *[Aunt Karin]*, verb and adverb *[finally found]*, critical noun (direct object) *[the toothbrush]*, prepositional phrase *[in the sink]*, and final (most often another prepositional phrase) *[of the bathroom]*. The first prepositional phrase (PP) suggested a horizontal orientation (*the toothbrush [in the sink] of the bathroom*) of the critical noun in 10 of these sentences and a vertical orientation (*the toothbrush [in the cup] beside the mirror*) in the other 10 sentences; half of both orientations matched, and half mismatched the orientation of the picture. The remaining 20 sentences were filler items; they did not contain words that were used earlier in other phases of the experiment. A total of 10 filler sentences were followed by a comprehension question, requiring a "yes" or "no" response, to ensure that participants would actually read the sentences.

Procedure

The three phases of the experiment, lasting for about 40 min in total, were presented as three different experiments, which lasted 10–15 min each. To test whether participants were unaware of the connections between the experiments, a post-experiment interview was conducted. Participants were asked whether anything stood out to them. If the answer was negative or not relevant, they were asked whether they could think of a connection between the experiments. Finally, it was revealed that words from the first phase were used in the third phase, and participants were asked whether they had noticed this.

During the word–picture verification phase of the experiment, participants saw a 700-ms presentation of a word, followed by a 200-ms presentation of a picture. They indicated as quickly as possible whether the picture matched the word. Each word–picture combination was shown twice in random order. Participants first received three practice trials. In the third phase, participants read sentences presented one by one on a computer screen, while their eye movements were recorded by a Tobii 2150 eye tracker with a 21-inch monitor and 50-Hz refresh rate. Participants were seated at approximately 65 cm from the computer screen, yielding a visual angle of 0.34° per character. The participants' head movement was not restrained. When the eye tracker indicated that the participant was looking

[1] The mental rotation task was initially intended to be used for the assessment of individual differences, but the variability within our sample was too small to make this type of analysis meaningful.

at the fixation point, a sentence was presented. After reading and understanding a sentence the participant pressed the space bar, after which a new fixation point appeared. After every five sentences, a true/false comprehension question was presented to make sure that participants were reading for comprehension. Participants were instructed to read as naturally as possible. The reading phase started with five practice sentences.

Results

Analyses of data of the word–picture verification task show that participants made highly accurate responses in deciding whether a picture matched the preceding word ($M = 96\%$ correct, range 86–99%). Furthermore, responses to the comprehension questions during the reading task were above chance ($M = 88\%$ correct, range 67–100%), which indicates that participants, as instructed, naturally read the sentences; they did not just guess but did not read the sentences too thoroughly either.

For analysis of the eye-tracker data, we assessed first-pass reading times and total reading times with a custom program, written in Visual Basic. A first-pass reading time is the sum of fixation times (fixations >80 ms on one fixation point) in a region of text from first entering the region until exiting the region; the total reading time is the sum of all fixations in a region, including regressive and forward fixations. A fixation < 80 ms as well as fixations ±3 standard deviations from a participant's condition mean were excluded from the analyses. All in all, less than 6% of the reading times were removed. Each sentence contained five areas of interest (AOIs): subject *[Aunt Karin]*, verb and adverb *[finally found]*, critical noun (direct object) *[the toothbrush]*, prepositional phrase *[in the sink]*, and final (most often another prepositional phrase) *[of the bathroom]*. AOIs were calculated by length of words and were programmed as a specific area within particular coordinates. The key area of interest was always the fourth one, the prepositional phrase that implied an orientation for the target object.

Analyses of variance (ANOVAs) were conducted to investigate effects of condition on each of the five regions on both first-pass and total reading times. Counterbalancing list was included as a between-subject factor because it interacted significantly with condition (Pollatsek & Well, 1995a, 1995b). Q2 However, effects for the list variable are not reported given the lack of theoretical relevance.

Figure 1 suggests that first-pass reading times in the mismatch condition are significantly longer than first-pass reading times in the match condition in the critical PP that constrained the orientation of the target object (e.g., *in the sink*), whereas reading times in other regions do not differ between conditions. ANOVAs confirm this impression; first-pass reading times differed significantly between the match and mismatch conditions, $F(1,\ 26) = 5.55$, $p < .03$, $MSE = 2{,}774$; $\eta^2_p = .18$. There was no significant effect in any of the other regions (all Fs < 1). For the total reading times, shown in Figure 2, the same qualitative pattern was found as that for the first-pass reading times. There was a marginal effect of condition on the fourth region, the first PP, $F(1,\ 26) = 4.09$, $p < .06$, $MSE = 8{,}520$; $\eta^2_p = .12$. There were no other significant differences (all Fs < 2.55, ps > .12). Thus, both first-pass and total reading times show that condition had an effect on the first PP of the sentences; reading times were longer for the mismatch condition than for the match condition; furthermore, condition had no effect on other regions.

Postexperiment interviews showed that none of the participants was aware of the purpose of the

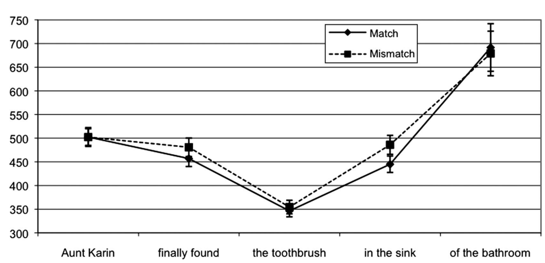

Figure 1. *Average first-pass reading times (in ms) per region. Error bars depict standard errors of the means.*

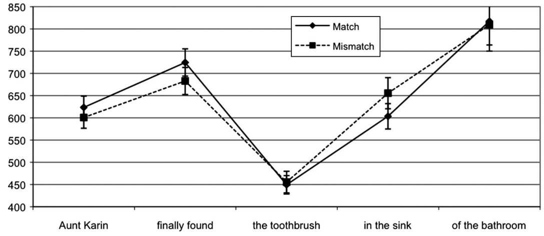

Figure 2. *Average total reading times (in ms) per region. Error bars depict standard errors of the means.*

first part in relation to the third part of the experiment. Although 3 participants noticed that certain objects they had seen in the first part were repeated in the third part of the experiment, none of the participants was aware of the manipulation of orientation.

Discussion

Prior incidental exposure to a picture of an object in a particular orientation affected reading times for phrases that described the location and thereby constrained the orientation of an object from the same category. For example, participants who had previously briefly seen a picture of a horizontal nail, read the phrase *into the wall* faster than participants who had seen a picture of a vertical nail. These results are in accordance with earlier results (Pecher et al., 2009; Stanfield & Zwaan, 2001) and extend these results by showing that not only does language affect the processing of visual information, but visual memory also influences language processing.

These results paint a picture of reading comprehension as a multimodal integration process that uses not only linguistic representations, but also sensory representations (at least if they were recently linked to individual words). Even though the input during reading is linguistic (in this case visually presented), a sensory representation that is consistent with the evolving representation leads to faster processing than a representation that is inconsistent with the evolving representation. Apparently, comprehension involves achieving coherence among activated mental representations irrespective of their modality. This conclusion is consistent with theories and models that assume that language comprehension involves the incremental and multimodal integration of linguistic and nonlinguistic information (Altmann & Mikovic, 2009; Q3 Elman, 2009; Barsalou, 1999; Zwaan, 2004). At the phenomenological level, these results provide a demonstration of the everyday situation of having a prior filmic experience—for example, having seen the *Lord of the Rings* movies colours a later reading experience of the novels on which they were based.

Given that the effect already occurred during first-pass reading times, it appears to be immediate, perhaps automatic. Furthermore, given that it materialized in the fourth region of the sentences, the region in which orientation is implied, it suggests that, just like motor resonance effects, they appear to occur at the location where they are within linguistic focus, which is consistent with the linguistic-focus hypothesis (Taylor & Zwaan, 2008). The visual memory paradigm Q4 described in this article allows for further investigation of the subtle interplay between linguistic input and visual memory. A question at the top of the research agenda should be what the exact role of each is in the comprehension process.

Original manuscript received 12 November 2009
Accepted revision received 29 March 2010
First published online day month year

REFERENCES

Altman, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*, 502–518. Q5

Altmann, G. T. M., & Mirkovic, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, *33*, 583–609. Q3

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

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105,* 14325–14329.

Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science*, *33,* 1–36.

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Glenberg, A. M. (1997). What memory is for. Q6 *Behavioral and Brain Sciences*, *20,* 1–55.

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

Hagoort, P., & Van Berkum, J. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *362,* 801–811.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, *42,* 335–346.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85,* 363–394.

Palmer, S. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Pecher, D., Van Dantzig, S., Zwaan, R. A., & Zeelenberg, R. (2009). Language comprehenders retain implied shape and orientation of objects. *The Quarterly Journal of Experimental Psychology*, *1,* 1–7.

Pollatsek, A., & Well, A. D. (1995a). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning,* Q2 *Memory, & Cognition*, *21,* 785–794.

Pollatsek, A., & Well, A. D. (1995b). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27,* 767–780. Q2

Pylyshyn, Z. W. (1986). *Computational cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.

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

Taylor & Zwaan (2008). Q4

Zwaan, R.A. (2004). The immersed experiencer: Q7 Toward an embodied theory of language comprehension. In B.H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 44, pp. 35–62). New York: Academic Press.

Zwaan, R. A., Adriaans, M., Pillay, D., & Aveyard, M. (2010). *Incidentally acquired visual experiences affect later reading*, Manuscript submitted for publication. Q1

Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic mental representations in language comprehension. *Cognitive Science*, *28,* 611–619. Q7

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123,* 162–185. Q8

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13,* 168–171. Q9

PQJE502579

Queries

Stephanie I. Wassenburg and Rolf A. Zwaan

Dear Author
Please address all the numbered queries on this page which are clearly identified on the proof for your convenience.

Thank you for your cooperation

Q1 Zwaan, R. A., Adriaans, M., Pillay, D., & Aveyard, M. (2010). "subitted" changed to 2010, year OK? [manuscripts in preparation or submitted need date: "Use the year of the draft you read (not 'in preparation') in the text citation" (*APA5*: **4.16,** p. 264)]

Q2 Pollatsek, A., & Well, A. D. (1995). Changed to Pollatsek & Well (1995a), (1995b), as it is really two references.

Q3 Altmann & Mirkovic (2009). Mirkovic in refs, Mikovic in text. Which is correct?

Q4 Taylor & Zwaan, 2008. Not in refs. Please provide full reference.

Q5 Altman, G. T. M., & Kamide, Y. (2007). Should it be Altmann, as in next ref.? Not cited in the text. Please add a citation or delete. [APA requires each entry in the reference list to be cited in the text (*APA5*: **4:01,** p. 215)]

Q6 Glenberg, A. M. (1997). Not cited in the text. Please add a citation or delete.

Q7 Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Not cited in the text. Please add a citation or delete.

Q8 Zwaan, R. A., & Radvansky, G. A. (1998). Not cited in the text. Please add a citation or delete.

Q9 Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Not cited in the text. Please add a citation or delete.