Acta Psychologica 145 (2014) 98–103

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/ locate/actpsy

Automatic perceptual simulation of first language meanings during second language sentence processing in bilinguals

Nikola Vukovic *, John N. Williams

Department of Theoretical and Applied Linguistics, University of Cambridge, UK

ARTICLE INFO

Article history:
Received 18 September 2013
Received in revised form 6 November 2013
Accepted 8 November 2013
Available online 10 December 2013

PsycINFO classification:
2720 - Linguistics & Language & Speech
2340 - Cognitive Processes

Keywords:
Embodied Cognition
Bilingual
Semantics

ABSTRACT

Research supports the claim that, when understanding language, people perform mental simulation using those parts of the brain which support sensation, action, and emotion. A major criticism of the findings quoted as evidence for embodied simulation, however, is that they could be a result of conscious image generation strategies. Here we exploit the well-known fact that bilinguals routinely and automatically activate both their languages during comprehension to test whether this automatic process is, in turn, modulated by embodied simulatory processes. Dutch participants heard English sentences containing interlingual homophones and implying specific distance relations, and had to subsequently respond to pictures of objects matching or mismatching this implied distance. Participants were significantly slower to reject critical items when their perceptual features matched said distance relationship. These results suggest that bilinguals not only activate task-irrelevant meanings of interlingual homophones, but also automatically simulate these meanings in a detailed perceptual fashion. Our study supports the claim that embodied simulation is not due to participants' conscious strategies, but is an automatic component of meaning construction.

1. Introduction

The nature of meaning – how we represent and transmit it – stands as one of the foundational problems in cognitive science. The ability to communicate using language is a distinctively human trait, yet the mechanisms describing how exactly the linguistic code becomes meaningful are much debated. Understanding a sentence, on traditional accounts, consists of computing a propositional representation which, in essence, specifies necessary and sufficient conditions needed for the sentential content to be true (e.g., Fodor, 1998). However, any semantic theory in which meaning is the exclusive domain of amodal computation runs into the serious problem of symbol grounding (Harnad, 1990; Searle, 1980). An alternative approach to this issue has recently been presented in theories whose aim is to explain the operation of the evolutionarily recent linguistic system as built upon, and grounded in, phylogenetically much older sensorimotor brain structures. On this embodied interpretation of cognition, language comprehension is achieved by recruiting the very same resources which are used for action, perception, and emotion. Linguistic meaning, in other words, arises through performing mental simulations of sentential content, during which we reactivate fragments of experience formed during past perception and action (Barsalou, 1999).

Evidence from behavioural and neuroimaging studies supports the claim that, when understanding language, people perform mental simulation, and that this simulation is embodied. Studies indicate that we mentally represent perceptual and visual information described in comprehended sentences (Glenberg & Kaschak, 2002; Horton & Rapp, 2003; Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002). For example, Zwaan et al. (2002) had participants read sentences implying certain actions (e.g., *The ranger saw the eagle in the sky*), and subsequently verify pictures that were congruent or incongruent with the implied event (e.g., a picture of an eagle with outstretched wings vs. an eagle in a nest). They found that people were significantly faster to respond to pictures that were consistent with the implied sentential content, suggesting that they were simulating the shape of objects and animals involved. Similarly, behavioural studies have confirmed that comprehenders simulate a range of other perceptual features, such as orientation (Stanfield & Zwaan, 2001), location (Bergen, Lindsay, Matlock, & Narayanan, 2007), visibility conditions (Horton & Rapp, 2003), and motion (Kaschak et al., 2005). Interestingly, mental simulation also includes movement direction (Glenberg & Kaschak, 2002), as well as other properties such as action duration. In Matlock (2004), for example, participants read the sentence "*Road 49 crosses the desert*" much slower if they were previously told that the desert was 400 miles in diameter, than when they were told that the desert was only 30 miles in diameter.

Neuroimaging studies also support the view that language comprehension crucially involves simulation of sensory, motor, and emotional content. Pulvermüller (2005) found that reading action words such as "kick" or "run" reliably activates areas in the motor cortex used for performing arm or leg movements. Similarly, auditory perception has overlapping neural correlates with auditory imagination (Halpern,

* Corresponding author at: Department of Theoretical and Applied Linguistics, English Faculty Building, University of Cambridge, 9 West Road, Cambridge CB3 9DP, UK.
E-mail address: nv254@cam.ac.uk (N. Vukovic).

0001-6918/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.actpsy.2013.11.002

Zatorre, Bouffard, & Johnson, 2004). Speer, Reynolds, Swallow, and Zacks (2009) found a congruent simulation pattern, not only during individual word presentation, but also during sentence or story reading. They conclude that understanding a story produces brain activation very similar to that during performing, imagining, or observing relevant actions/events in the world.

In sum, then, both behavioural and neuroimaging data suggest that language comprehension is more than the kind of purely symbolic process assumed by traditional theories. Comprehension seems to involve activating sensorimotor neural resources to perform detailed, multimodal, and dynamic simulation of perceptual and motor states, through which the linguistic codes we use in communication become meaningful.

The studies outlined above have sparked an active discussion in the field, highlighting the need for additional inquiry into the automaticity and universality of perceptual simulation processes, especially across diverse linguistic populations. Concerns have been raised that the reported findings might not reflect a process that is basic to language comprehension, but is rather the result of participants' conscious decision to imagine a described scene after they have already understood the meaning (Kiefer & Pulvermüller, 2012; Mahon & Caramazza, 2009). The aim of the present study was to address this question by exploiting a known fact about bilingual lexical processing — automatic, and unconscious, activation of lexical representations in the non-current language (usually the L1 during an L2 processing task).

Previous research has demonstrated that bilinguals cannot switch off one of their languages during comprehension (De Bruijn, Dijkstra, Chwilla, & Schriefers, 2001; de Groot, Delmaar, & Lupker, 2000; De Groot & Nas, 1991; Schwartz & Kroll, 2006; Van Assche, Duyck, & Hartsuiker, 2012; Van Heuven, Dijkstra, & Grainger, 1998). This non-selective activation of both L1 and L2 words at the same time has been found to persist irrespective of task demands and participants' cognitive strategies (Dijkstra, Timmermans, & Schriefers, 2000; Dijkstra & Van Hell, 2003). For example, research in bilingual language processing has demonstrated that bilinguals non-selectively activate both meanings of a homophone word (e.g., Lagrou, Hartsuiker, & Duyck, 2011). These task-irrelevant L1 meanings become active not only in single word reading, but also in a sentential and semantic context (Lagrou, Hartsuiker, & Duyck, 2012). In conclusion then, language non-selective and automatic lexical access is a robust finding confirmed by decades of research in bilingual language processing.

Our study aims to exploit this known fact about L2 lexical processing to investigate a crucial issue in the literature: the automaticity of embodied simulatory processes in language comprehension. The bilingual mind seems to be the perfect testing ground for this hypothesis derived from monolingual research. We know, as previously described, that people perform detailed perceptual simulation during language processing; however, it is debated whether such simulation is consciously initiated and maintained, or if it proceeds automatically during comprehension.

Here, we adapt the experimental design used in Winter and Bergen (2012), which was originally introduced by Stanfield and Zwaan (2001). Winter and Bergen (2012) had participants read sentences about objects, where some sentences implied visual distance (*You are looking at the microphone on the other side of the stage*) and some implied closeness (*You are looking at the microphone in your hand*). Afterwards, participants were presented with pictures of mentioned objects which, crucially, varied in size. Their task was to make a judgment on whether the object in the picture was mentioned in the previous sentence. The reasoning was that if readers mentally simulate perceptual features, including distance, then they should be faster to verify images congruent with sentence-implied distance. Indeed, participants were found to respond faster to small images of mentioned objects after reading sentences implying visual distance, and to big images following sentences implying closeness, suggesting that they visually simulated the sentence meaning. No such effect was obtained in control trials, when participants responded to pictures of objects not mentioned in the sentences.

In the present study the above paradigm was modified so as to test the main hypothesis on a bilingual population and, exploiting the well-known fact about non-selective bilingual language processing, ascertain more closely the nature of embodied simulation processes. Dutch-English bilinguals heard sentences in their L2 (English), after which they saw images which varied in size, and had to make a judgment on whether the depicted objects were mentioned in the previously heard sentences. In the Winter and Bergen (2012) study, participants would read sentences on the screen and would manually indicate when they have read them, thus advancing to the next trial. This procedure may be problematic, because it gives the participants the freedom and time for slow (re-)reading, and might possibly favour strategies wherein participants purposefully imagine the sentence content in anticipation of the picture. Because of this, we decided to use auditory presentation of sentences, which is a more rapid procedure which minimises variation in stimulus presentation time. Additionally, Winter and Bergen (2012) used sentences within which the position of the target object-word varied. Here, all critical targets were object-words located uniformly across sentences (always in sentence-final position). Most importantly, however, the current experiment is more likely to uncover automatic processes because of a novel modification of the classic design: all the critical object-words were interlingual English-Dutch homophones (words which sound similar in both languages, but denote something different in each).

If, following second language research, bilinguals unconsciously activate L1 meanings when processing L2 sentences containing interlingual homophones, and if it is true, as Embodied Cognition suggests, that meaning activation includes performing perceptual simulation, then we should see evidence of simulation processes even on task-irrelevant L1 meanings of words. Alternatively, if perceptual simulation is a process initiated intentionally by participants, then we might expect that unconsciously activated semantic material will not participate in a top-down driven simulation. If the former account is true, we would expect to see slower latencies in the matching homophone condition (where image size matches sentence implied distance), and faster ones in the mismatching condition (image size does not match sentence implied distance). The reasoning behind this prediction is as follows. Winter and Bergen (2012) saw facilitation of congruent responses in monolinguals when the pictures matched the implied sentence distance. Here, however, participants have to reject, not accept, the critical pictures, because they show task irrelevant representations of L1 homophone meaning. For example, after hearing the English sentence *"On the plate in front of you, you can see a bone,"* participants would see a picture showing beans — the word for which in Dutch is "boon"/bo:n/. The picture shown varied in size between conditions, such that it could match or mismatch the implied sentence distance (for an example of sentence–picture pairs, see Table 2). Critically, if bilinguals end up perceptually simulating both the task-relevant L2 meaning and the irrelevant L1 meaning, it should be harder for them to reject the critical pictures. This is because, even though task-irrelevant, the pictures would match the participants' mental simulation on one additional dimension — visual distance. The main RT differences to look out for, then, will be those between congruent and incongruent homophone target pictures.

The results of the current experiment could have important implications for debates both in Embodied Cognition and bilingual processing. As we have seen, Embodied Cognition theory predicts mental simulation of meaning that is automatic and central to language comprehension. Crucially, on this account, this simulation would be one that is multimodal, and embodying a specific perspective — namely, one that mirrors actual perception. Therefore, understanding sentences in our experiment should lead to simulation of objects at specific distances, and whose size is consistent with the first-person perspective of an immersed observer (Barsalou, 2002). If we find evidence of such