Brief article

Evidence for the activation of sensorimotor information during visual word recognition: The body–object interaction effect

Paul D. Siakaluk a,*, Penny M. Pexman b, Laura Aguilera a, William J. Owen a, Christopher R. Sears b

a Department of Psychology, University of Northern British Columbia, Prince George, BC, Canada V2N 4Z9
b Department of Psychology, University of Calgary, Calgary, AB, Canada T2N 1N4

Received 5 July 2006; revised 15 November 2006; accepted 10 December 2006

Abstract

We examined the effects of sensorimotor experience in two visual word recognition tasks. Body–object interaction (BOI) ratings were collected for a large set of words. These ratings assess perceptions of the ease with which a human body can physically interact with a word’s referent. A set of high BOI words (e.g., mask) and a set of low BOI words (e.g., ship) were created, matched on imageability and concreteness. Facilitatory BOI effects were observed in lexical decision and phonological lexical decision tasks: responses were faster for high BOI words than for low BOI words. We discuss how our findings may be accounted for by (a) semantic feedback within the visual word recognition system, and (b) an embodied view of cognition (e.g., Barsalou’s perceptual symbol systems theory), which proposes that semantic knowledge is grounded in sensorimotor interactions with the environment.

Keywords: Embodied cognition; Lexical processing; Visual word recognition

* Corresponding author. Tel.: +1 250 960 6120; fax: +1 250 960 5744.
E-mail address: siakaluk@unbc.ca (P.D. Siakaluk).

0010-0277/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
1. Introduction

What do humans know and how is this knowledge acquired? How is this knowledge accessed when recognizing an object, understanding a concept, or comprehending a story? Traditionally, cognitive scientists sought to answer these questions principally by analyzing internal cognitive processes, such as the manner in which symbolic representations are manipulated via rules (e.g., Cowart, 2004). Although this approach is indisputably important, much of the research it has generated has either ignored or down-played the importance of our sensorimotor interactions with the environment. A growing field within cognitive science, known as embodied cognition, examines how sensorimotor interactions with the environment are integral in the acquisition of knowledge and to the development of cognitive processes that bear on that knowledge (Barsalou, 1999; Clark, 1997; Cowart, 2004; Lakoff & Johnson, 1999; Pecher & Zwaan, 2005; Wilson, 2002).

According to Wilson (2002), the embodied cognition framework incorporates six distinct claims regarding the manner in which the mind, the body, and the external world interrelate to help organisms successfully cope with environmental demands. Two of these claims are directly relevant to the research presented in this paper. The first is that “(c)ognition is for action. The function of the mind is to guide action, and cognitive mechanisms such as perception and memory must be understood in terms of their ultimate contribution to situation-appropriate behavior” (p. 626). The second claim is that “off-line cognition is body based. Even when decoupled from the environment, the activity of the mind is grounded in mechanisms that evolved for interaction with the environment – that is, mechanisms of sensory processing and motor control” (p. 626).

Our research examined whether sensorimotor information gained through body–object interaction is relevant in a task as abstract and removed from motor activity as visual word identification. If such sensorimotor information can be shown to influence visual word recognition, then this result would provide support for the claims of the embodied cognition perspective outlined above. As Pecher and Zwaan (2005) argued, “it is crucial for the embodied framework to demonstrate that cognition is grounded in bodily interactions with the environment... thus, it needs to be shown that sensorimotor patterns are activated when concepts are accessed.” (p. 3). Our research was motivated by this theoretical issue.

In her discussion of the claim that off-line cognition is body based, Wilson (2002) suggested that many on-line cognitive processes that originally evolved to help an organism directly deal with its environment may have been co-opted for different, less direct functions. These cognitive processes may also be used off-line, to run mental simulations of possible ways of interacting with the environment, avoiding the need to actually carry out each possible action. Recent studies have investigated these cognitive processes using one of two general approaches. First, researchers have used methodologies that employ relevant sensorimotor activities, just prior to or at the time of testing, that are proposed to be integral to the required responses. For example, experiments have shown that specific arm movements (e.g., pulling or pushing movements) influence responses when items have a particular valence (i.e., positive or
negative) (Markman & Brendl, 2005), or are congruent with a particular implied action during sentence comprehension (Glenberg & Kaschak, 2002). This type of approach has also been applied in virtual reality environments. In one such study, Christou and Bülthoff (1999) reported superior recognition of scenes when participants had been directed to actively (vs. passively) explore those scenes in a virtual reality environment. Similarly James et al. (2002) reported faster recognition of objects following active rotation (vs. passive viewing of object rotations). As these examples demonstrate, sensorimotor activities, such as arm movements or active exploration of virtual environments, facilitate cognitive processes underlying responding.

The second approach is to examine sensorimotor effects using methodologies that do not employ relevant sensorimotor activities prior to or at the time of testing. For example, Lakoff and Johnson (1999) provided detailed accounts of the ways in which abstract metaphors are understood through knowledge gained via sensorimotor interactions with the world. As an example, consider one particularly vivid metaphor, Bad Is Stinky, as in “This movie stinks” (p. 50; emphasis in the original). This metaphor is comprehensible because of prior sensorimotor experience with foul-smelling objects (e.g., rotting food) and the desire to avoid them. A second example of this approach (although seldom discussed as embodied cognition effects) is the research examining the effects of imageability and concreteness in visual word recognition tasks (e.g., Cortese, Simpson, & Woolsey, 1997; James, 1975; Kroll & Mervis, 1986; Strain, Patterson, & Seidenberg, 1995). (Because imageability and concreteness are conceptually similar and highly correlated, we will consider them jointly). A great deal of research has shown that words with referents that can be more easily imaged/sensed have richer mental representations that enable them to be recognized more rapidly in word recognition tasks. Imageability/concreteness effects have also been reported in word association tasks (de Groot, 1989), and definition tasks (Sadoski, Kealy, Goetz, & Paivio, 1997).

Although it is difficult to separate precisely the effects of sensory experience and motor experience on cognitive processing, we argue that imageability/concreteness ratings primarily gauge sensory experience. If this is the case, then an important and interesting question is whether variables that capture primarily motor experience also influence conceptual processing. A few recent studies have demonstrated that motor information, as gauged by variables measuring object manipulability (Magnie, Besson, Poncet, & Dolisi, 2003; Myung, Blumstein, & Sedivy, 2006; Wolk, Coslett, & Glosser, 2005), influences object recognition. To our knowledge, only one study has examined the effects of motor information on word recognition. Myung et al. (2006) examined the effects of motor information on priming using an auditory lexical decision task. Related primes and targets shared manipulation features. Manipulation features were defined as “general actions on an object that involve body movements and typically are associated with its intended usage” (p. 228). For example, keys and screwdrivers share manipulation features at a general action pattern level because they require similar wrist movements. Myung et al. reported a significant priming effect: target words preceded by a word that shared common manipulation features (e.g., key-screwdriver) were responded to faster than target words preceded by a word...
that did not share common manipulation features (e.g., bar-screwdriver). These results suggest that sensorimotor information was activated in a priming paradigm and facilitated auditory lexical decisions.

1.1. The present research

The present research builds on the Myung et al. (2006) findings. We examined the effects of sensorimotor knowledge on recognition of individual words (we did not use a priming paradigm). That is, we examined whether sensorimotor knowledge influences word recognition, even when that knowledge is not preactivated by related primes.

To create the stimuli for our study, we asked a group of undergraduate students to rate words as to the ease or difficulty with which a human body can physically interact with each word's referent. These ratings allowed us to extract a variable we call body–object interaction (BOI). We based this rating scheme on Cortese and Fugett's (2004) rating scheme for collecting imageability ratings. They had participants rate words “as to the ease or difficulty with which they arouse mental images” (p. 387). After collecting the ratings we then selected items such that the set of high BOI words were matched to the set of low BOI words for both imageability and concreteness (as well as several other lexical and semantic variables). Thus, any effects of BOI should be attributable to sensorimotor information that has been gained through prior bodily interactions with the objects.

We assumed that effects of sensorimotor experience (BOI effects) could be explained in terms of feedback activation from semantics to orthography and to phonology. The notion of feedback activation in the visual word recognition system has recently been invoked to explain a number of effects (Hino & Lupker, 1996; Hino, Lupker, & Pexman, 2002; Pecher, 2001; Pexman & Lupker, 1999; Stone, Vanhoy, & Van Orden, 1997; Strain et al., 1995). The feedback activation account has the following important assumptions. First, the system is fully interconnected, with feedforward and feedback connections between sets of orthographic, phonological, and semantic units (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Plaut, McClelland, Seidenberg, & Patterson, 1996). Second, the influence of semantic feedback is determined by the type of feedback connections between the semantic units and either the orthographic units or the phonological units. That is, to the extent that the mappings between semantic and other sets of units are many-to-one, there will be a facilitatory influence of semantic feedback (relative to the situation where the same connections are one-to-one). Third, lexical decision task (LDT; is it a word?) responses are based primarily on the activation of the orthographic units and phonological lexical decision task (PLDT; does it sound like a word?) (and naming) responses are based primarily on the activation of the phonological units.

Within this framework, we predicted that effects of BOI would be facilitatory. The rationale for this prediction comes from a number of previous research findings, all of which suggest that semantic richness (e.g., high imageability/concreteness, high
number of semantic features, high number of word senses) tends to facilitate word identification (James, 1975; Pexman, Lupker, & Hino, 2002; Rodd, Gaskell, & Marslen-Wilson, 2002). To the extent that high BOI words also have richer semantic representations, they should generate stronger feedback activation from semantics to both orthography and to phonology, and thus they should be identified more rapidly than low BOI words.

We examined the effects of BOI using a LDT, to assess the influence of BOI on feedback activation from semantics to orthography, and using a PLDT, to assess the influence of BOI on feedback activation from semantics to phonology. Typically, semantic effects are larger in the LDT if the task is made more difficult by including pseudohomophones (e.g., *brane*) as foils (e.g., James, 1975; Pexman & Lupker, 1999). We therefore included pseudohomophone foils in the LDT to make the task maximally sensitive to BOI effects. In this context, the prediction we derived from the embodied cognition perspective was that high BOI words would be identified faster than low BOI words.

2. Method

2.1. Participants

Sixty undergraduate students from the University of Northern British Columbia participated in the experiments for bonus course credit: 30 participants each in the LDT and in the PLDT. All participants were native English speakers and reported normal or corrected-to-normal vision.

A separate group of 25 participants rated the stimuli for body–object interaction, and another group of 25 participants rated the stimuli for number of features. These participants were drawn from the same population as the individuals who participated in the LDT and PLDT experiments, and also received bonus course credit.

2.2. Stimuli

An initial pool of 234 words was first created. Each word had only one entry in the *ITP Nelson Canadian Dictionary* (1997) and all had noun definitions listed first. The words were randomly ordered in two questionnaires. For the body–object interaction questionnaire, participants were instructed to rate each word according to the ease or difficulty with which a human body can physically interact with the word’s referent. For the number of features questionnaire, participants were instructed to rate each word according to the number of features possessed by the word’s referent (the instructions for this questionnaire were those used by Toglia & Battig (1978)). For the body–object interaction questionnaire a scale from 1 to 7 was placed to the right of each word, with 1 indicating low body–object interaction and 7 indicating high body–object interaction. For the number of features questionnaire the same scale was placed next to each word, with 1 indicating low number of features and 7
indicating high number of features. Participants were encouraged to use the entire scale in making their ratings, and were instructed that all the words were nouns and to base their ratings on this fact.

Based on these ratings, 48 words were selected for use in the experimental tasks (hereafter the experimental items). Twenty-four of the words were rated as being high in body–object interaction (e.g., mask) and the other 24 words were rated as being low in body–object interaction (e.g., ship). These two word groups were matched for length, printed frequency, subjective familiarity, number of orthographic neighbours, phonological feedback inconsistency, number of features, number of senses, number of associates, semantic distance, and, importantly, both imageability and concreteness (all $p > .15$). The descriptive statistics for the experimental words are presented in Table 1.

Forty-eight pseudohomophones with extant word bodies (e.g., brane), matched to the experimental items on length, were used in both the LDT and the PLDT. In addition, 96 nonwords were used in the PLDT. The stimuli are listed in Appendix A.

### 2.3. Apparatus and procedure

For both the LDT and the PLDT the stimuli were presented on a colour VGA monitor driven by a Pentium-class microcomputer running DirectRT software (2006). A trial was initiated by a fixation marker that appeared at the center of the computer display. The fixation marker was presented for 1000 ms and was then replaced by a letter string. For the LDT, participants were instructed to decide whether each item was a real English word or not. For the PLDT, participants were instructed to decide whether each item sounded like a real English word or not. For both tasks, ‘yes’ responses were made by pressing the “?” key and ‘no’ responses were made by pressing the “z” key on the computer keyboard. Participants were instructed to respond as quickly and as accurately as possible. Response latencies were measured to the nearest ms. Stimulus order was randomized separately for each participant. The intertrial interval was 2000 ms.

For the LDT, each participant first completed 20 practice trials, consisting of 10 words similar in normative frequency to the experimental items and 10 pseudohomophones. For the PLDT, each participant first completed 20 practice trials, consisting of 5 words similar in normative frequency to the experimental items, 5 pseudohomophones, and 10 nonwords.
3. Results

Response latencies less than 250 ms or more than 2500 ms were treated as outliers and removed from the dataset. In addition, for each participant, response latencies greater than 2.5 SD from the cell mean of each condition were also treated as outliers and removed. A total of 51 observations (3.54% of the data) were removed from the LDT and 50 observations (3.47% of the data) were removed from the PLDT. The mean response latencies of correct responses and mean error percentages for all stimuli are presented in Table 2. In the subject analyses ($F_1$) BOI was a within-subject manipulation; in the item analyses ($F_2$) BOI was a between-item manipulation.

3.1. LDT

In the analysis of the response latency data, the 35 ms BOI effect was statistically significant, $F_1(1, 29) = 14.39$, MSe = 1278.89, $p < .005$, $\eta^2 = .33$; $F_2(1,46) = 5.45$, MSe = 2558.94, $p < .05$, $\eta^2 = .10$. Responses to high BOI words were faster than responses to low BOI words. There was no BOI effect on error rates (both $F$s < 1).

3.2. PLDT

In the analysis of the response latency data, the 36 ms BOI effect was statistically significant, $F_1(1, 29) = 12.80$, MSe = 1485.47, $p < .005$, $\eta^2 = .30$; $F_2(1,46) = 7.35$, MSe = 1964.13, $p < .01$, $\eta^2 = .13$. Like the lexical decision responses, phonological lexical decisions were faster to high BOI words than to low BOI words. Error rates were not analyzed because there were too few errors for a proper analysis (less than 1% overall).

Table 2
Mean response latencies (in milliseconds) and standard errors, and mean response error percentages and standard errors

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>LDT</th>
<th>PLDT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SE</td>
</tr>
<tr>
<td><strong>Response latencies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High BOI</td>
<td>628</td>
<td>17.1</td>
</tr>
<tr>
<td>Low BOI</td>
<td>663</td>
<td>21.8</td>
</tr>
<tr>
<td>BOI effect</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>Pseudohomophones</td>
<td>817</td>
<td>30.3</td>
</tr>
<tr>
<td>Nonwords</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>Response errors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High BOI</td>
<td>3.36</td>
<td>0.8</td>
</tr>
<tr>
<td>Low BOI</td>
<td>3.42</td>
<td>0.8</td>
</tr>
<tr>
<td>BOI effect</td>
<td>0.06</td>
<td>–</td>
</tr>
<tr>
<td>Pseudohomophones</td>
<td>6.08</td>
<td>0.9</td>
</tr>
<tr>
<td>Nonwords</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
4. Discussion

The purpose of this study was to determine whether sensorimotor knowledge, as measured by BOI, influences the processing of written words. We observed a facilitatory BOI effect – faster responses to high BOI words than to low BOI words – in both the LDT and the PLDT.¹ These results, like the results of Myung et al. (2006; Experiment 1), suggest that lexical semantics (conceptual knowledge that is accessed in word recognition tasks) includes information about sensorimotor experience with objects. This requires an expansion of the traditional notion of lexical semantic content, to include information gained through nonverbal, primarily motor, experience.

The semantic feedback framework, introduced in Section 1.1, provides an explanation for the mechanism by which BOI facilitates visual word recognition. According to this framework, words with relatively richer semantic representations (measured by BOI in the present study) provide stronger feedback activation to orthography and to phonology, allowing faster responding. What has to be explained, of course, is how sensorimotor information is incorporated in semantic memory. On this issue, an embodied cognition framework, such as that provided by Barsalou’s (1999) perceptual symbol systems theory, provides one account for how sensorimotor knowledge, as measured by BOI, may be stored in and retrieved from memory.

Barsalou (1999, 2003a, 2003b; Barsalou, Simmons, Barbey, & Wilson, 2003) proposed the perceptual symbol systems theory to describe how semantic knowledge is grounded in sensorimotor experience. According to the theory, knowledge is acquired through sensorimotor experience and retrieval of that knowledge involves simulation or partial reenactment of the sensorimotor states implicated at encoding. Semantic knowledge is thus represented in terms of simulators. Activation for a particular concept involves activating a subset of the knowledge that is represented in a simulator and running a simulation that reenacts some of what is known about that concept. As Barsalou (2003a) put it,

“Consider the category of CARS. Visual information about how cars look is integrated with auditory information about how they sound, olfactory information about how they smell, motor information about driving them, somatosensory information about feeling the ride in them, and emotional information associated with speed, dangerous situations, etc. The resulting representation is a distributed system… that establishes knowledge about CARS.” (p. 1180)

Thus, sensorimotor information is proposed to be incorporated as conceptual knowledge. As such, the BOI effects we have observed are quite compatible with perceptual symbol systems theory.

¹ The high BOI words and the low BOI words were matched as closely as possible on all of the lexical and semantic variables listed in Table 1, but there were still small differences between the two sets of words on a few of these variables. Statistically equating the two sets of words on each of these variables via analyses of covariance (ANCOVA) demonstrated that these differences were not important. The effect of BOI was statistically significant in all of the ANCOVAs with one minor exception: when concreteness was the covariate in the analysis of the LDT response latencies the BOI effect was marginally significant (p = .06).
4.1. Conclusion

In the present work, we report a new semantic effect: the BOI effect. This effect underscores the need for theories of visual word recognition to include sensorimotor knowledge as an essential form of knowledge contained in lexical semantics. Our results provide support for the idea that off-line cognition is to a significant extent embodied – that is, influenced by prior sensorimotor experience with the world.

Acknowledgements

This research was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) discovery grants to P.D.S., P.M.P., and C.R.S. We thank Shirley Fortin, Kim Foster, Keri Locheed, and Linda Kerswell for their help in data collection. We also thank three anonymous reviewers for their very helpful comments.

Appendix A. Items used in the experiments

A.1. High BOI words

belt brick couch crown crumb dish drum fence flute gift grape lamp mask pear pipe purse rope skirt stool suit tape thorn tool vest.

A.2. Low BOI words

cake cliff cloud clown creek dirt ditch dorm flame flood juice kite lace leaf mist pond seed shelf ship silk smog torch tribe tube.

A.3. Pseudohomophones

berd boal boan bote brane crain dait doar drane gaim gard goast gote group gurl hoam hoap hoze jale jerm jirk klaim koast nale noat nurve rane rong rore roze scail sheat shurt skalp skarf sleap smoak stawl stoar swet teath thret tode treet tutch werk wheet.

A.4. Nonwords

bame beal besh bime binch bope bram brame brank brate bulch chate cheen clace clirp cron crong cruss dack dake dawk dreeb drench duss fage fill fitch flane flang flef flet foom full fung gake gick glank gless grabe grafe gurse hain hape hean helt hife hine jick jote kine koose loke ludge meep merch moch nent nerte pake pame pape pell petch pilk pleap poote potch pribe prog pung rame rask rell scuff scug shate shink slirt soat spale spen spoop stort strup tain talt tane tark thurn tinch toin trake treen trine turt vank yelf.
References


