Acta Psychologica 156 (2015) 98–103

Contents lists available at ScienceDirect

Acta Psychologica

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

Time course of action representations evoked during sentence comprehension

Alison W. Heard, Michael E.J. Masson *, Daniel N. Bub

Department of Psychology, University of Victoria, PO Box 1700 STN CSC, Victoria, BC V8W 2Y2, Canada

ARTICLE INFO

Article history:
Received 11 May 2013
Received in revised form 10 January 2014
Accepted 30 January 2014
Available online 18 February 2014

PsycINFO classification:
2340 (cognitive processes)

Keywords:
Action representations
Eye movements
Imagery
Mental simulation
Motor resonance
Sentence comprehension

ABSTRACT

The nature of hand-action representations evoked during language comprehension was investigated using a variant of the visual–world paradigm in which eye fixations were monitored while subjects viewed a screen displaying four hand postures and listened to sentences describing an actor using or lifting a manipulable object. Displayed postures were related to either a functional (using) or volumetric (lifting) interaction with an object that matched or did not match the object mentioned in the sentence. Subjects were instructed to select the hand posture that matched the action described in the sentence. Even before the manipulable object was mentioned in the sentence, some sentence contexts allowed subjects to infer the object's identity and the type of action performed with it and eye fixations immediately favored the corresponding hand posture. This effect was assumed to be the result of ongoing motor or perceptual imagery in which the action described in the sentence was mentally simulated. In addition, the hand posture related to the manipulable object mentioned in a sentence, but not related to the described action (e.g., a writing posture in the context of a sentence that describes lifting, but not using, a pencil), was favored over other hand postures not related to the object. This effect was attributed to motor resonance arising from conceptual processing of the manipulable object, without regard to the remainder of the sentence context.

1. Introduction

Substantial evidence supports the proposal that mental representations of actions are evoked during language comprehension (e.g., Glenberg & Kaschak, 2002; Kaschak & Borreggine, 2008; Zwaan & Taylor, 2006). These representations may support an active, deliberate mental simulation of motor activity described in a sentence, which in turn could contribute to successful comprehension. In addition, an action representation may become active when an object is mentioned simply by virtue of its inclusion in the general conceptual knowledge associated with that object (e.g., Bub & Masson, 2012; Masson, Bub, & Lavelle, 2013). This elicitation of action representations may be quite separate from any ongoing, overtly constructed mental simulation and may result merely from identifying or holding in working memory an object concept. We report an experiment that provides evidence consistent with this distinction between two different roles played by action representations during comprehension.

Both neuroimaging and behavioral evidence support the view that action representations are evoked during language comprehension when the message conveys information about manipulable objects. Research using fMRI has shown that when listening to action-based verbs, or nouns representing manipulable objects, activation of somatotopically relevant areas of motor cortex occurs (Rueschemeyer, Brass, & Friederici, 2007; Tettamanti et al., 2005). Activations of this kind are also found during actual performance of actions described by sentences (Hauk, Johnsrude, & Pulvermüller, 2004). This evidence indicates that during the processing of words that denote manipulable objects or actions, the motor cortex creates an embodied representation that may then become reactivated upon reencountering that event. Behavioral evidence for a relationship between language comprehension and embodied action representations includes demonstrations that reach and grasp responses can be made more quickly when cued in the context of sentences that describe related as opposed to unrelated actions (e.g., Bub & Masson, 2010; Masson, Bub, & Newton-Taylor, 2008; Masson, Bub, & Warren, 2008; Masson et al., 2013).

It has been suggested that embodied representations of action are a form of mental simulation, and that it is possible to differentiate two distinct types of mental simulation, which we will refer to as *motor imagery* and *motor resonance* (Barsalou, 2008; Kent & Lamberts, 2008; Moulton & Kosslyn, 2009). Moreover, there is neuroimaging evidence that supports a dissociation between these two mechanisms (Willems, Toni, Hagoort, & Casasanto, 2010). Motor imagery is the explicit construction of mental representations of action in working memory and these representations may be maintained indefinitely (Grezes & Decety, 2001; Jeannerod, 1994). We intend this term to include the possibility that instead of motor-based imagery a mental simulation may involve perceptual imagery representing visible aspects of action (e.g., a mental image of a moving hand). In contrast, motor resonance

* Corresponding author. Tel.: +1 250 721 7536; fax: +1 250 721 8929.
E-mail addresses: allyh@uvic.ca (A.W. Heard), mmasson@uvic.ca (M.E.J. Masson), dbub@uvic.ca (D.N. Bub).

0001-6918/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.actpsy.2014.01.017

is an automatic recruitment of an action representation that usually is temporally localized to the presentation of an inducing stimulus rather than extending over time periods typical of motor imagery (Masson et al., 2013; Zwaan & Taylor, 2006). For example, Zwaan and Taylor had subjects read sentences one segment at a time, with segments advancing as subjects rotated a knob. The required rotation was either clockwise or counterclockwise and that direction was either consistent or inconsistent with the action implied by the verb in the sentence (e.g., closing or opening a bottle). Reading time was influenced by the consistency between the verb and the direction of rotation, but this influence was restricted to the sentence frame containing the verb and did not extend beyond that. Similarly, Masson et al. cued subjects to perform a reach and grasp response while listening to a sentence that described an action that was congruent or incongruent with respect to the cued action. Congruency effects on action responses were obtained when the action cue was presented during or shortly after mention of a manipulable object in the sentence. When the action was cued later in the sentence, the congruency effect had dissipated. Masson et al. also showed, however, that if subjects were induced to engage in motor imagery during presentation of the sentence (by requiring subjects to be prepared to pantomime the action described in the sentence), the congruency effect on action performance was sustained through to the end of the sentence.

An additional finding in one of the motor imagery experiments reported by Masson et al. (2013, Exp. 6) was that a cued action relevant only to the manipulable object mentioned in a sentence, but not to the particular action being described, was primed along with the action that fully conformed to the sentence context. For instance, given the sentence

(1) *To clear the shelf, Jack lifted the pen*

the action that is congruent with the full context is a horizontally oriented precision grip. That action was primed by the sentence context, but so, too, was the action corresponding to using a pen (a writing posture), even though writing was not implied by the sentence context as a whole, only by the specific manipulable object mentioned in the sentence. Masson et al. proposed that priming of the lifting action was associated with overt mental simulation of the action described in the sentence context, and that priming of the functional action (writing in this example) resulted from motor resonance induced by mention of the manipulable object.

In the experiment reported here, we sought to examine this proposed distinction using a new method that was intended to provide a fine-grained assessment of the time course of these two mechanisms for invoking action representations. Our approach is an adaptation of the visual–world paradigm, which uses tracking of eye movements and fixations to provide a real-time assessment of the knowledge representations that are active during language processing. Research using this paradigm most often analyzes the likelihood with which a subject fixates a particular image or object over time when that item is presented as one member of an array of stimuli (e.g., Huettig, Rommers, & Meyer, 2011). Past research has shown that subjects' eyes tend to fixate on stimuli that are congruent with current mental operations, providing researchers with information about the time course of mental processes during language comprehension (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

With this method, we have a means of probing the status of action representations without requiring subjects to execute reach and grasp actions. Rather than cuing subjects to perform hand actions at various points during the presentation of a sentence context, we had them view a display of four hand postures. Two of the postures were relevant to the manipulable object mentioned in the context sentence, one representing what we refer to as a *volumetric* action (used to move or pick up an object), and the other corresponding to a *functional* action (performed when using the object for its intended purpose). The sentence described either a volumetric or a functional interaction with the object, so only one of these two postures fits the full sentence context. For example, the following sentence fits the functional action of pressing buttons rather than the volumetric action of lifting the device:

(2) *Bobby used the cellphone to text his friends.*

The other two postures in a display depicted the functional and volumetric actions associated with an object not mentioned in the sentence.

To encourage the generation of overt mental simulation of the activity described by sentence contexts, subjects were instructed to select the depicted hand posture that matched the action described by the sentence context. We assumed that if motor representations corresponding to an action afforded by the manipulable object mentioned in the sentence were evoked, then the subject's eyes would tend to move to the displayed posture that depicted that action. Under the instruction to select the hand posture that fits the sentence context, eye fixations should eventually be heavily concentrated on the relevant posture. The eye-tracking method we used allows us to ask when, during the course of sentence processing, particular action representations are elicited. For example, constructing a mental simulation of a described action may be postponed until the target object is mentioned. Alternatively, mental simulation may begin as soon as sentence information provides some constraints on the set of possible actions that fit with what has been presented so far. Consider the following two sentences, in which the action's distal goal (outlining a pattern) is presented early versus late:

(3) *To outline her pattern, Grace used the marker.*
(4) *Grace used the marker to outline her pattern.*

When the distal goal is mentioned first, as in (3), some constraints on the possible relevant actions are established in the first clause of the sentence and construction of a mental simulation may begin at a relatively early stage during sentence comprehension. But when mention of the distal goal is moved to the end of the sentence, as in (4), few if any constraints on the relevant action are present until the verb or perhaps even the manipulable object is mentioned.

Of particular interest in this experiment is the possible role played by motor resonance. Masson et al. (2013) demonstrated that an action representation relevant to a manipulable object mentioned in a context sentence was primed, even though it was not consistent with the specific action described by the sentence. Mention of an object by itself seems capable of evoking both functional and volumetric action representations (e.g., Bub & Masson, 2012). This type of activation appears to be a product of motor resonance elicited as part of the general knowledge associated with a manipulable object. Either functional or volumetric action representations may be elicited through this mechanism when the context consists only of the name of an object (Bub & Masson, 2012), but it is not clear whether this principle holds when a complete sentence context specifying a particular action is presented. Masson et al. (Exp. 6) showed that functional action representations were elicited through motor resonance when sentence contexts described volumetric interactions with objects and subjects were induced to overtly simulate those interactions, but they did not test the converse possibility (volumetric actions elicited when subjects apply motor imagery to a functional action). Moreover, their method required subjects to execute overt reach and grasp actions. These response demands may have influenced the pattern of motor activations that were observed. In the present experiment, we sought evidence for evocation of action representations when subjects were not explicitly making reach and grasp actions. It is important to note that the evidence for motor resonance that we anticipate finding would mean that an action representation inconsistent with the general sentence context, but congruent with the manipulable object mentioned in the sentence, would be evoked and that this would happen at the same time as the fully congruent action representation was highly active as part of an overt mental simulation of the activity described by the sentence context. The onset of motor resonance defined this way should occur during or shortly after mention of the object noun. It may persist into the interval following that mention because efforts to maintain the object concept in working memory while generating a mental image of the action implied by the sentence could sustain activation of the object concept, including its associated action representations—even those not related to the current mental simulation. Masson et al. (Exp. 6) did not probe any points of the sentence following mention of the object noun so the possible persistence of motor resonance under mental imagery instructions is an open question.

2. Method

2.1. Subjects

Thirty-two students at the University of Victoria were tested. All subjects were either native or fluent speakers of English and received extra credit in an undergraduate psychology course for their participation.

2.2. Materials

Three functional and three volumetric hand actions were selected and grouped as three action pairs containing one action of each type. For each pair of actions, four objects were identified that were deemed by the authors to be associated with those actions. The action pairs and associated object names are shown in Table 1. Photographs of a male hand formed into postures depicting the six actions were taken and rendered as digital gray-scale images. These images were used to create displays containing four hands arranged in a 2×2 matrix and displayed against a white background. Each display included two of the three functional/volumetric pairs listed in Table 1. Each of the three combinations of two pairs was used equally often, and for each combination, four different displays were created by varying which hand image appeared in which of the four spatial locations. Across the full set of displays, each hand image appeared equally often in each of the four spatial locations of the display. Left- and right-handed versions of each display were generated for use with left- and right-handed subjects, respectively. Each display measured 12.0° of visual angle both horizontally and vertically when presented on the monitor viewed by the subject. The visual angle from the center of one hand to the center of either adjacent hand was 5.9°.

A set of 288 sentences was constructed, half describing a functional interaction with an object and the other half describing a volumetric interaction. All sentences were structured so that an actor was mentioned first, followed by a functional (e.g., *used*) or volumetric (e.g., *lifted*) verb, the name of a manipulable object, and finally a clause elaborating the purpose of the action. A second version of each sentence was constructed by moving the final clause (distal goal) to the beginning of the sentence, as in the following examples.

(5) *The little girl picked up the crayon to show it to her friend.*
(6) *To show it to her friend the little girl picked up the crayon.*

The assignment of sentences to the distal-goal-first or distal-goal-second version was counterbalanced across subjects so that each sentence was presented equally often in the two versions. Each subject was tested on 72 sentences in each of the four conditions defined by the factorial combination of sentence type (volumetric or functional) and position of distal goal (first or second clause). In addition, two different assignments of hand displays to sentences were used (one for each half of the subjects). We ensured that for all assignments, one of the two functional–volumetric posture pairs in a display was congruent with the manipulable object mentioned in the associated sentence. An additional set of 24 sentences was constructed for use as practice items.

Each of the 12 objects listed in Table 1 was used in 12 different sentences of each interaction type (functional or volumetric). The sentences were recorded by a female speaker whose first language was English. During the experiment, four different pseudorandom assignments of the hand displays to the auditory sentences were used across subjects, with the constraint that the display chosen for a sentence included the pair of hand images associated with the manipulable object mentioned in the sentence. For example, for a sentence about someone lifting a pencil, the required pair of hand images would be the pencil grip and the horizontal pinch and the other two images could be either of the two remaining pairs (see Fig. 1). The subject's tasks were to listen to a sentence while simultaneously monitoring the display of hand images and to select the hand that matched the action depicted by the sentence. Selection was indicated by having the subject move the computer mouse and point and click it over the target image.

2.3. Procedure

Subjects were tested in a quiet room equipped with an SR Research EyeLink 1000 tower-mount eye-tracking system. Eye position was recorded from one eye 1000 times per second, although the subject viewed the display binocularly. The tower mount provided a chin and forehead rest to stabilize head position. A computer mouse was positioned in front of the subject who used his or her dominant hand to move and click it to make a response. Stimuli were displayed to the subject on a cathode ray tube monitor placed 78 cm from the subject. Display of visual stimuli and the auditory presentation of sentences through external speakers was controlled by a Macintosh Intel G3 computer. A Dell computer was used to record eye-fixation data.

Calibration of the subjects' gaze was completed at the beginning of the experiment, as well as after any movement during rest breaks. This was achieved by having the subject fixate on nine points arranged in a 3×3 grid extending across the viewing area of the monitor. After the initial calibration 24 practice trials were completed and then experimental trials began. Subjects were instructed to listen to the sentences for comprehension and simultaneously to monitor the display of four hands in order to select which hand image matched the action described in the sentence. The 288 critical trials were presented in a random order and a rest break was offered after every block of 60 trials. Each trial began with a fixation cross at the center of the screen for at least 500 ms. Once the subject fixated the cross, it was replaced by the hand display. Onset of the auditory sentence followed 500 ms later.

2.4. Data analysis

A program for editing digital audio files (Sound Studio, version 4.5.4) was used to define the temporal boundaries for the segments into which each sentence was divided. For sentences structured with the distal goal in the second position, the segments were the subject phrase, verb phrase, object, and final clause (distal goal) (e.g., *Jen/used the/pen/to write to her friend*). The same segments were used for the distal-goal-first sentences, except that the distal-goal phrase was placed at the beginning of the sentence. Because these sentences ended with the manipulable object it was possible that moving the mouse to click on the target hand image could occur after that noun had been enunciated. Therefore, we defined an additional period of 1000 ms after the end of these sentences as a final sentence segment. Fixations made during that interval, up to the time of the mouse-click response or the end of the interval, whichever came first, were included in our analyses and

Table 1
Hand actions and associated object names used in the experiments.

Functional action	Volumetric action	Associated object names
Pencil grip	Horizontal pinch	Crayon, marker, pen, pencil
Spray grip	Vertical grasp	Hair spray, insect spray, room spray, spray paint
Thumb grip	Horizontal grasp	Cellphone, Gameboy, iPod, television remote

are reported as a separate sentence segment for the distal-goal-first sentences.

Using the eye-tracking data gathered from the EyeLink 1000 system, we determined the proportion of time within each of the sentence segments that a subject spent viewing each of the four hands in the visual display. Adjacent rectangular areas of equal size were used to define the regions associated with the displayed hands and any fixation falling within a hand's region was included in the total for that hand. For each sentence segment, we computed the proportion of that segment's audio playing time that was spent on each hand. Eye fixations that occurred between the onset of the sentence and the subject's mouse-click response were included in the analyses we report. The relevant hand conditions were defined by the factorial combination of action type (functional or volumetric) and relation to the object mentioned in the sentence context (congruent or incongruent).

Because the subjects were free to respond at any time, we had no control over where within a particular sentence the sequence of included fixations would end. Nevertheless, for all of the conditions of the design, every subject had measurable fixation time in each of the sentence segments for at least one sentence. Our results are reported as the proportion of fixation time within a sentence segment that was spent on each of the four available hand images.

3. Results

Subjects were highly accurate ($M = 96.2\%$) when selecting the hand image that matched the sentence context. Eye-fixation data from trials where a response error occurred were excluded from our analysis. The mean proportion of time spent fixating each type of hand posture as a function of sentence context and sentence segment is shown in Figs. 2 and 3 for distal-goal-first and distal-goal-second sentences, respectively. For both sentence structures and for both sentence types (volumetric and functional), subjects clearly spent longer fixating the target hand posture, particularly as more of the sentence context was presented. This impression was verified by analyses of variance (ANOVA) conducted separately for the distal-goal-first and distal-goal-second sentences. Separate analyses were conducted because these two classes of sentence were segmented differently. In each case, the independent variables were sentence type (volumetric, functional), sentence segment, hand type (volumetric, functional), and hand congruency (congruent, incongruent), all of which were manipulated within subjects. Preferential viewing of the hand posture specifically relevant to the full sentence context (both the manipulable object and the verb) developing over time would be revealed as a four-way interaction between all of these factors. Both ANOVAs produced highly significant four-way interactions, $F(4, 124) > 500$, $p < .001$, in both cases.

Inspection of Fig. 2 shows that when the distal goal appeared as the first clause, subjects began to favor the target hand posture very early—even in the first sentence segment in the case of functional sentences, $F(1, 93) = 85.99$, $p < .001$ for the comparison between the target posture and the other three postures. For volumetric sentences, the two volumetric hand postures were favored by the time the sentence's second segment was reached [$F(1, 93) = 100.11$, $p < .001$, for the contrast between volumetric and functional actions] and thereafter the target volumetric action was favored. These early effects emerged before the verb or the manipulable object were mentioned, indicating that subjects used subtle clues (including knowledge about the features of the set of sentences used in the experiment) to deduce what actions were most likely to be described by later segments of the context sentence. For example, the first clause in a sentence such as

(7) *To outline her pattern Grace used the marker*

could readily lead subjects to anticipate that the relevant hand posture would be a writing grip.

Considering now the sentences in which the distal goal was the second clause, we can see in Fig. 3 that for both volumetric and functional

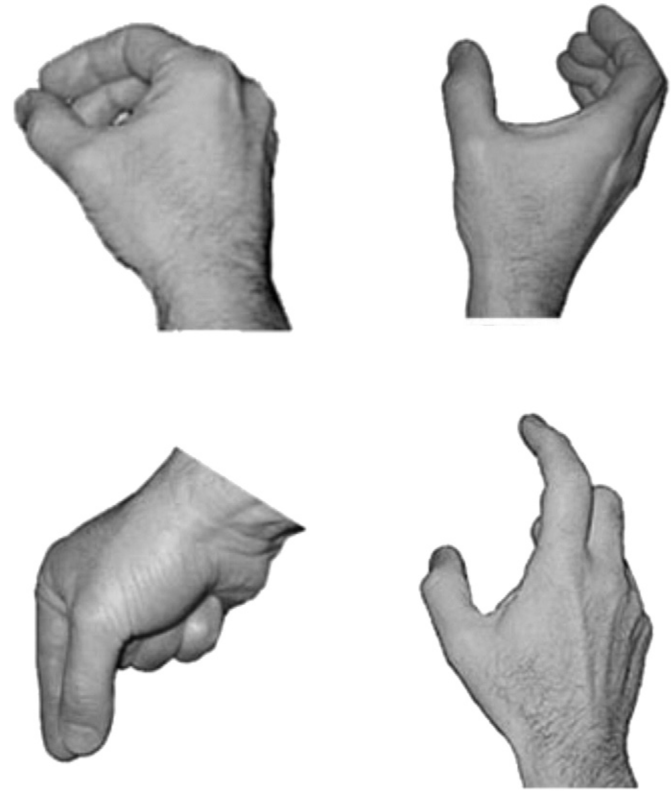

Fig. 1. An example display of four hand images.

sentences, eye fixations favored the two hand images relevant to the verb used in the sentence (used vs. lifted) as soon as the verb was mentioned, $F(1, 93) = 31.77$, $p < .001$ and $F(1, 93) = 13.64$, $p < .001$ for volumetric and functional sentences, respectively, for the contrast between verb compatible and verb incompatible hand images. This was the first point at which the sentences conveyed information that offered any distinction at all between the four hand images being displayed. As soon as the object was mentioned, the hand relevant to the combination of the already-mentioned verb and the object was strongly favored, $F(1, 93) = 436.73$, $p < .001$ and $F(1, 93) = 652.70$, $p < .001$ for volumetric and functional sentences, respectively, for the contrast between the target hand image and the other three hand images combined.

In addition to evidence for early construction of an overt mental simulation during sentence comprehension, Figs. 2 and 3 provide clear evidence for motor resonance driven only by the manipulable object mentioned in the sentence context. For all four sentence conditions (defined by position of the distal goal and the type of verb), the hand posture that was relevant only to the manipulable object, but not to the verb, received a greater share of fixation time than did the other two irrelevant hand postures. For example, with sentence (7) as the context, the horizontal precision posture would be congruent with the object (but not the verb) whereas any of the actions not related to a marker would be incongruent. Importantly, this congruency effect emerged only as the object was being mentioned or shortly thereafter, in the final sentence segment. The one exception to this constraint was for functional sentences in the distal-goal-first condition (see Fig. 2), where the volumetric action congruent to the object noun showed an advantage over the other two incongruent hand images in the second sentence segment. But this advantage disappeared in the next two segments and re-emerged only in the final segment. The advantage in the final sentence segment for the object-congruent posture over the two incongruent postures averaged together was shown to be statistically significant in an ANOVA that also included sentence

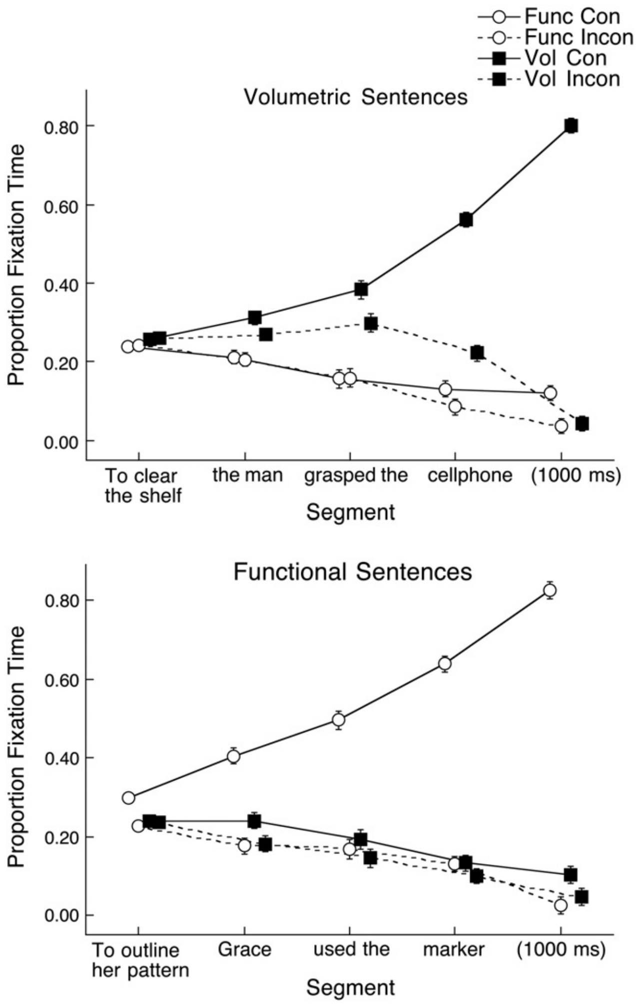

Fig. 2. Mean proportion of fixation time associated with each hand-image condition as a function of sentence type (volumetric or functional) and sentence segment. These data are for sentences in which the distal goal was the first clause in the sentence. Data points within each segment are staggered for visual clarity. Error bars are 95% within-subject confidence intervals computed separately for each sentence segment and are appropriate for comparing means within each segment (Loftus & Masson, 1994; Masson & Loftus, 2003). Some error bars are smaller than the symbols denoting the conditions.

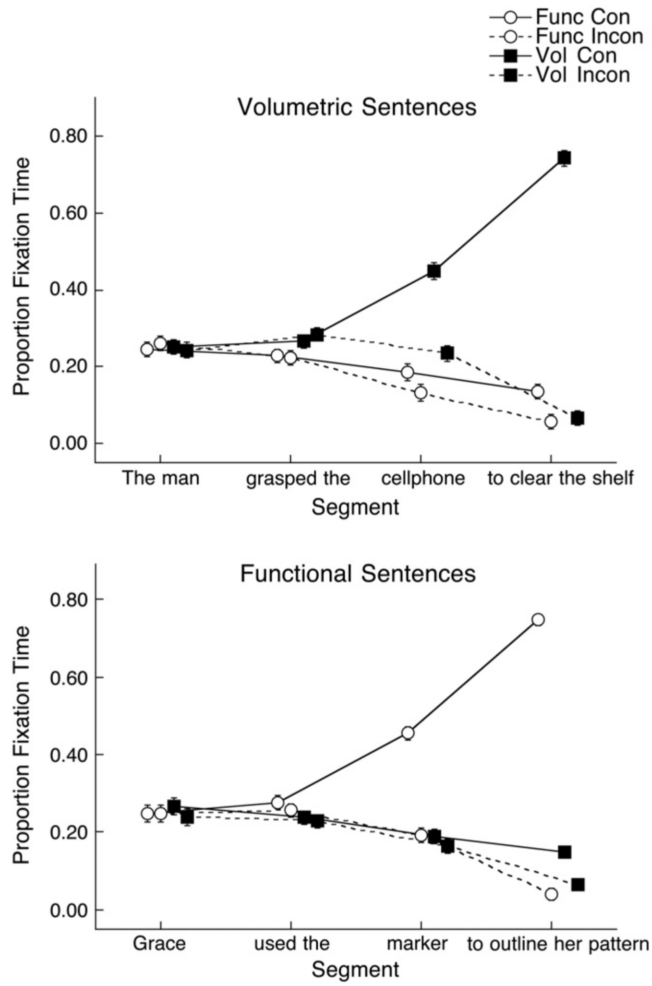

Fig. 3. Mean proportion of fixation time associated with each hand-image condition as a function of sentence type (volumetric or functional) and sentence segment. These data are for sentences in which the distal goal was the second clause in the sentence. Data points within each segment are staggered for visual clarity. Error bars are 95% within-subject confidence intervals computed separately for each sentence segment and are appropriate for comparing means within each segment. Some error bars are smaller than the symbols denoting the conditions.

structure and sentence type as factors. This analysis yielded a main effect for posture congruency, $F(1, 31) = 362.10$, $p < .001$, with a greater proportion of viewing time associated with the object-congruent posture relative to the incongruent postures (.13 vs. .05). There was also a three-way interaction involving the posture comparison and the two sentence variables, $F(1, 31) = 8.84$, $p < .01$, indicating that the congruency effect varied somewhat across sentence conditions, but the effect was clearly present in all cases (mean differences ranged from .07 to .10). The fact that the advantage for the object-congruent posture was maintained even as the subject made an overt selection of a different posture (the one congruent with both the object and verb) makes this evidence for motor resonance particularly impressive. The effect clearly cannot be a direct result of the motor imagery underlying the overt mental simulation of the action described by the sentence because that action is different from the posture that is relevant only to the object.

4. Discussion

Our goal was to investigate the involvement of action representations in language comprehension. The eye-tracking methodology we used allowed us to generate a detailed time course for the evocation of hand-action representations while subjects listened to sentences describing interactions with manipulable objects. Unlike previous studies that used action-priming tasks (e.g., Bub & Masson, 2010; Masson, Bub, & Newton-Taylor, 2008; Masson, Bub, & Warren, 2008; Masson et al., 2013), our experiment did not require object-related motor responses. Responses of that kind could have elicited action representations that otherwise might not be evoked in the course of sentence comprehension. The purpose of the experiment was to determine the time course of the evocation of hand-action representations as well as the mechanisms underlying that evocation. Two distinct mechanisms were considered: (a) mental simulation involving motor imagery and (b) motor resonance.

Requiring subjects to overtly select a hand image that was consistent with the action described in a context sentence produced eye fixations that were rapidly attracted to the target image. Indeed, as soon as the sentence made available information that allowed some level of discrimination among the hand images being displayed, subjects favored the implicated image(s). Deeper progression into the sentences led to the one fully relevant target image being strongly favored. This progression is consistent with the view that subjects began to construct an overt mental simulation of the action described in the sentence as soon as any constraining information was available. In our experiment, this result was obtained without requiring subjects to plan or make reach and grasp responses that coincided with actions described in the context sentences.

In addition to evidence for mental simulation constructed through motor imagery, we also obtained evidence for a form of motor resonance elicited specifically by the manipulable object rather than by the combined meaning of the object and verb. The displayed hand posture that was associated with the object, but not the verb, mentioned in the context sentence showed a small advantage in the share of fixation time over the two unrelated hand images. Masson et al. (2013, Exp. 6) obtained evidence for a similar phenomenon when subjects were induced to engage in motor imagery by the requirement to pantomime the volumetric action described by a context sentence. Under those conditions, both volumetric and functional actions associated with the critical object were primed to a similar degree. Priming of the functional action could not be attributed to motor imagery because the ongoing mental simulation involved an action (lifting) that would be in opposition to the functional action (using) associated with the manipulable object. Instead, functional action representations were deemed to have been elicited by motor resonance arising from conceptual processing of the manipulable object. In the present experiment, the object-congruent action that also fits the verb context received the large majority of fixation time, and the other object-congruent action was only weakly preferred over the fully irrelevant actions. This discrepancy between the two experiments is most likely due to the requirement here for subjects to use visual guidance in placing the mouse cursor on the target hand image.

We suggest that the evocation of action representations associated with a manipulable object, but not supported by the verb context, arises from the fact that mental simulation of the described action must entail conceptual processing of the critical object. That processing is assumed to evoke both functional and volumetric action representations associated with the object, regardless of sentence context. But this is an implicit form of mental simulation (motor resonance) and not a direct product of motor imagery (Masson et al., 2013). Indeed, by requiring subjects to imagine the hand action corresponding to the sentence context, they must construct a representation (e.g., functional action) that would be in opposition to the other action (e.g., volumetric) related to the mentioned object. Consistent with this suggestion, eye fixations begin to favor the object's alternate action over at least one of the two unrelated actions at the point where the type of object can first be discerned and this advantage continues into the next (final) segment of the sentence. For example, in the top panels of Figs. 2 and 3, which depict the results for volumetric sentences, the critical object cannot be inferred until it is mentioned directly and it is not until that point that fixation on the object-congruent functional action exceeds fixation on the object-incongruent functional action. In the case where the distal goal appears in the first clause of the sentence, and a functional action is described, it is sometimes possible to infer the manipulable object before it is mentioned and we see in Fig. 2 (lower panel) evidence for motor resonance even before the verb or manipulable object is mentioned. In this case, once the action representation is evoked, it is sustained all the way through the sentence, in parallel with continuous mental imaging of the critical object.

Unlike the temporally limited indications of motor resonance seen in the results reported by Masson et al. and Zwaan and Taylor (2006), here we have an example of sustained motor resonance. We suggest that in this case the activation persists because of active maintenance of the inferred manipulable object in working memory as an explicit mental simulation is constructed (see Zwaan & Taylor, 2006, p. 8, for a similar proposal). Indeed, with functional sentences where the distal goal is presented first (lower panel of Fig. 2), subjects appear able to infer at least on some trials the relevant object just on the basis of the distal goal. Mental simulation can begin at that point and motor resonance associated with the inferred manipulable object follows shortly thereafter.

4.1. Conclusions

The results from this experiment provide a distinction between two proposed mechanisms for eliciting motor representations during language comprehension: motor resonance and motor imagery. The basis of this distinction, activation of object-specific action representations under motor resonance and activation of action representations relevant to full sentence context through motor imagery, is consistent with the analogous distinction reported by Masson et al. (2013). An important contribution of the present experiment is that subjects were not required to make reach and grasp actions while comprehending sentences, so it is unlikely that motor planning could have contributed to the observed pattern of effects. Instead, evocation of action representations appears to be a natural consequence of comprehending sentences that describe interactions with manipulable objects.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada through discovery grants awarded to M. Masson and to D. Bub. We thank Kendall Foster and Marnie Jedynak for their assistance in preparing and conducting the experiment.

References

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.
Bub, D. N., & Masson, M. E. J. (2010). On the nature of hand-action representations evoked during written sentence comprehension. *Cognition*, *116*, 394–408.
Bub, D. N., & Masson, M. E. J. (2012). On the dynamics of action representations evoked by names of manipulable objects. *Journal of Experimental Psychology: General*, *141*, 502–517.
Glenberg, A.M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*, 558–565.
Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping*, *12*, 1–19.
Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action word in human motor and premotor cortex. *Neuron*, *41*, 301–307.
Huettig, F., Rommers, J., & Meyer, A. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*, 151–171.
Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, *17*, 187–245.
Kaschak, M. P., & Borreggine, K. L. (2008). Temporal dynamics of the action-sentence compatibility effect. *Quarterly Journal of Experimental Psychology*, *61*, 883–895.
Kent, C., & Lamberts, K. (2008). The encoding–retrieval relationship: Retrieval as mental simulation. *Trends in Cognitive Sciences*, *12*, 92–98.
Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476–490.
Masson, M. E. J., Bub, D. N., & Lavelle, H. (2013). Dynamic evocation of hand action representations during sentence comprehension. *Journal of Experimental Psychology: General*, *142*, 742–762.
Masson, M. E. J., Bub, D. N., & Newton-Taylor, M. (2008). Language-based access to gestural components of conceptual knowledge. *Quarterly Journal of Experimental Psychology*, *71*, 869–882.
Masson, M. E. J., Bub, D. N., & Warren, C. M. (2008). Kicking calculators: Contribution of embodied representations to sentence comprehension. *Journal of Memory and Language*, *59*, 256–265.
Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, *57*, 203–220.
Moulton, S. T., & Kosslyn, S. M. (2009). Imagining predictions: Mental imagery as mental emulation. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *364*, 1273–1280.
Rueschemeyer, S., Brass, M., & Friederici, A. (2007). Comprehending prehending: Neural correlates of processing verbs with motor stems. *Journal of Cognitive Neuroscience*, *19*, 855–865.
Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science*, *268*, 1632–1634.
Tettamanti, M., Buccino, G., Saccuman, M., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, *17*, 273–281.
Willems, R. M., Toni, I., Hagoort, P., & Casasanto, D. (2010). Neural dissociations between action verb understanding and motor imagery. *Journal of Cognitive Neuroscience*, *22*, 2387–2400.
Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*, 1–11.