COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 41 (2017) 1613–1628

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12449

Embodied Simulations Are Modulated by Sentential Perspective

Wessel O. van Dam, Rutvik H. Desai

Department of Psychology, University of South Carolina

Received 13 January 2016; received in revised form 13 July 2016; accepted 19 July 2016

Abstract

There is considerable evidence that language comprehenders derive lexical-semantic meaning by mentally simulating perceptual and motor attributes of described events. However, the nature of these simulations—including the level of detail that is incorporated and contexts under which simulations occur—is not well understood. Here, we examine the effects of first- versus third-person perspective on mental simulations during sentence comprehension. First-person sentences describing physical transfer towards or away from the body (e.g., "*You threw the microphone*," "*You caught the microphone*") modulated response latencies when responses were made along a front-back axis, consistent with the action-sentence compatibility effect (ACE). This effect was not observed for third-person sentences ("*He threw the microphone*," "*He caught the microphone*"). The ACE was observed when making responses along a left-right axis for third-person, but not first-person sentences. Abstract sentences (e.g., "*He heard the message*") did not show an ACE along either axis. These results show that perspective is a detail that is simulated during action sentence comprehension, and that motoric activations are flexible and affected by the pronominal perspective used in the sentence.

Keywords: Grounded cognition; Semantics; Action sentence processing; Language

1. Introduction

Grounded theories of cognition propose that brain areas subserving perception and action also play a role in conceptual processing. According to such a view, real-world experiences stored in modality-specific sensory-motor areas form the basis for deriving lexical-semantic meaning (Barsalou, 1999, 2008; Glenberg, 1997; Pulvermüller, 1999, 2005). This theoretical position is fundamentally different from a symbolic (disembodied)

Correspondence should be sent to Rutvik H. Desai, Department of Psychology, University of South Carolina, 220 Discovery I, 915 Greene Street, 29205 Columbia, SC. E-mail: rutvik@sc.edu

view, which suggests that conceptual representations are symbolic and amodal (i.e., devoid of perceptual and action information), and that lexical-semantic meaning is derived purely through interactions between abstract symbols (Fodor, 1983; Katz & Fodor, 1963; Kintsch, 2008; Pylyshyn, 1984).

A number of behavioral studies have provided evidence that conceptual processing shares computational processes with perception and action (for a review see Fischer & Zwaan, 2008; Zwaan & Kaschak, 2008). For example, Tucker and Ellis (1998) showed that subjects were faster to make orientation judgments on manipulable objects (e.g., *tea pot*, *frying pan*) if the position of the object's handle protruded to the same direction as the manual response. This led the authors to conclude that the sensory-motor information related to the action that a manipulable object affords is automatically activated upon visual presentation of these objects. Similar compatibility effects have been shown for objects that afforded different types of grasps rather than effector-specific object affordances (Bub, Masson, & Cree, 2008; Masson, Bub, & Breuer, 2011). Tucker and Ellis (2001) showed that participants were faster to perform either a whole hand grasp or a precision grip, if the action was compatible with the grip afforded by the object (small vs. large object) they were presented with. Similarly, Helbig, Steinwender, Graf, and Kiefer (2010) showed that videos of a hand performing a certain action (e.g., a squeezing action) primed naming of objects that afford a congruent action (e.g., *pliers*) as compared with an incongruent action (e.g., *typewriter*). These findings provide evidence for a tight association between sensory-motor processes and conceptual representations accessed through pictures.

A number of studies have corroborated the findings of these studies by showing a similar involvement of sensory-motor processes in processing of concepts accessed through language. For example, Glenberg and Kaschak (2002) presented subjects with linguistic phrases describing a transfer of concrete objects towards (e.g., "*open the drawer*") or away from the body (e.g., "*close the drawer*"). They found that reaction times to sentences were modulated by whether the response (either moving towards or away from the body) was congruent with the directionality implied in the linguistic phrase. The authors labeled this congruency effect the action-sentence compatibility effect (ACE) and it has been taken as evidence that to understand a sentence like "*open the drawer*" comprehenders activate motoric information in a mental simulation of the described event. Likewise, Rueschemeyer, Lindemann, Van Rooij, Van Dam, and Bekkering (2010) have shown that comprehension of words denoting manipulable objects (e.g., *cup*, *hammer*) is facilitated by the prior planning of an action congruent to the prototypical use of the objects. In a similar vein, it has been shown that comprehenders mentally simulate perceptual attributes of objects (e.g., orientation, shape, color) during language comprehension (Connell, 2007; Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002). Brain imaging and patient studies provide similar lines of evidence (for reviews, see Kiefer & Pulvermüller, 2012; Binder & Desai, 2011; Meteyard, Cuadrado, Bahrami, & Vigliocco, 2012). These findings provide evidence that language comprehension entails the activation of experiential traces that are stored in modality-specific sensory motor brain regions.

Recent studies have shown that the types of information that are included in comprehenders' mental representations depend on contextual constraints. For example, Van Dam, Rueschemeyer, Lindemann, and Bekkering (2010) showed that compatibility effects between language and action depend on the type of conceptual properties emphasized by the linguistic context. They found compatibility effects for words that were preceded by a prime word that provided a context that emphasized the dominant action feature of the referent object (e.g., *thirst*—***cup***), but not when preceded by a prime word that emphasized a non-dominant action feature (e.g., *sink*—***cup***). In an imaging study, Desai, Conant, Binder, Park, and Seidenberg (2013) demonstrated that activation to the same action verb is modulated depending on whether the verb is used in literal action, metaphoric, or idiomatic sentences, such that idiomatic phrases resulted in the least activation of the motor system. Likewise, Rueschemeyer, Brass, and Friederici (2007) showed a similar neural response within the motor system between morphologically complex verbs built on motor stems (e.g., *begreifen = to comprehend*) and morphologically complex verbs built on abstract stems (e.g., *bedenken = to consider*). In the first case, the complex verb is a prefixed form of the simple motor verb *greifen* (to grasp), whereas in the latter case it is a prefixed form of the abstract verb *denken* (to think). These findings provide evidence that motor responses rather than being automatic and invariable, depend on the morphological context in which action words are encountered. These findings suggest that the activation of meaning attributes of words is a flexible process that depends on task or linguistic context.

A crucial question now concerns the precise content of motor activations and mental simulations. Here, we examine whether the perspective adopted in a sentence can affect the content of putative simulations and motoric activations. Numerous studies have provided evidence that perspective taking plays a crucial role in the comprehension of words and sentences (Black, Turner, & Bower, 1979; Morrow & Clark, 1988; Pustejowsky, 1995; Sanford & Garrod, 1998). For example, a study by Morrow and Clark (1988) showed that in a sentence like "*The mouse approached the fence,*" participants judged both the distance between agent and fence and the distance from where the situation was perceived to be smaller than in a sentence like "*The tractor approached the fence.*" These findings suggest that objects and agents denoted in a linguistic phrase can impose a perspective that is subsequently used in a reader's mental simulation or imagery. Furthermore, Brunyé, Ditman, Mahoney, Augustyn, and Taylor (2009) showed that pronouns can influence the perspective readers adopt in their mental simulations of described events. Participants in their study were presented with images taken from a first-person or third-person perspective for which they had to verify whether it matched or mismatched a description that made use of one of three pronouns (*I*, *You*, *He*). They found that for images in the first-person perspective participants' response times were lower if the description used the pronoun *I* or *You*, whereas for images in the third-person perspective participants' response times were lower if the description used the pronoun *He*. Furthermore, they found that when participants read an extended discourse that revealed the character identity, they were faster to respond to images in the third-person perspective if the description used the pronoun *I*. In a similar vein, it has been shown that the perspective imposed on an object by the linguistic phrase can modulate the perspective that comprehenders adopt. For example, Borghi, Glenberg, and Kaschak (2004) showed that after reading inside perspective sentences (e.g., "*You are driving a car*") participants were faster to verify inside parts of object images, whereas after reading outside perspective sentences (e.g., "*You are fueling a car*") participants were faster to verify outside parts of images. Likewise, while reading narratives that do not encourage the formation of mental images, comprehenders seem to adopt different perspectives in their mental simulations depending on the pronouns that were used in the action statements (Ditman, Brunyé, Mahoney, & Taylor, 2010). These findings suggest that the perspectives that comprehenders adopt in their mental simulation of perceptual attributes of objects and events are modulated by differences in the form of utterances (e.g., pronoun use) and by the perspective imposed on the object by the linguistic phrase.

These studies show that sentential perspective affects implicit visual imagery induced in service of sentence comprehension. An important question that remains is whether the motoric information is similarly modulated by differences in the perspective adopted in the sentence. One possibility is that action verbs like *throw* invoke a fixed partially abstracted simulation, regardless of the sentence context, and contain no perspective information. Alternatively, simulations could contain perspective information that is fixed (e.g., always first-person). Lastly, mental simulations may change based on the perspective described in the sentence. We investigated this issue by using a paradigm that has previously shown to elicit an ACE (Glenberg & Kaschak, 2002), which is the effect of facilitated sentence comprehension by prior planning of an action congruent to the movement direction implied in the sentence. We are interested whether changing the sentence perspective from a first-person (*You*) to a third-person perspective (*He*) influences observed congruency effects. To this end, we measured response times to sentences that described an abstract/concrete transfer of an object towards or away from the body. We hypothesized that the effect of movement preparation on sentence processing would interact with the sentence perspective and the orientation of the response device. Specifically, for concrete sentences that describe a transfer of an object towards or away from the body from a first-person perspective (e.g., "*You threw the microphone*"), we expect lower reaction times (RTs) in case the prepared motor response (*towards or away*) is congruent with the movement implied in the sentence. In the case of concrete sentences that describe the same transfer of an object but now from a third-person perspective (e.g., "*He threw the microphone*"), we do not expect an interaction with a prepared motor response *towards* or *away* from the body. Schwarzkopf, Weldle, Müller, and Konieczny (2011) failed to obtain an ACE for sentences with a third-person agent and patient (e.g., *The spectator pushes/pulls the pedestrian*). They argued that the perspective inherent in the verb (*pushing* is a movement away from the body and *pulling* is a movement towards the body) might not be sufficient to trigger an effective simulation. Alternatively, it could be argued that the perspective triggered by a verb depends on whether sentences use first-person or third-person agents and patients. That is, the same verb "*push*" might trigger a *towards-away* simulation when paired with a first-person agent (*I/You pushed*), whereas it might trigger a *left-right* simulation when paired with a third-person agent (*He pushed*).

A simulation along the *left-right* axis might be expected in which the agent is placed in the left side of conceptual space. Several studies have provided evidence that conceptual ordering of agents and patients follows the orthographic direction of the language (Chatterjee, Southwood, & Basilico, 1999; Maas & Russo, 2003). Therefore, in sentences from the third-person perspective (e.g., *He threw the microphone*) we expect participants to run a mental simulation of the action in which the agent occupies the left side of conceptual space (i.e., compatibility in terms of *towards* and *away* is now expected along the *left*-to-*right* axis).

Furthermore, we are interested in whether similar effects can be observed for sentences that describe an abstract transfer towards or away from the body (e.g., *He heard the message*). A number of studies have argued that abstract language similarly engages sensorimotor brain regions (Glenberg & Kaschak, 2002; Kiefer & Pulvermüller, 2012), whereas findings from other studies have questioned such an involvement in abstract language processing (Raposo, Moss, Stamatakis, & Tyler, 2009; Rueschemeyer et al., 2007).

2. Methods

2.1. Participants

Seventy-three people participated in the study. The study involved a between-subject manipulation of orientation of the response device. Responses were either made along a front-to-back axis (Experiment A) or left-to-right axis (Experiment B). Participants were between 18 and 34 years of age (M = 20.0, SD = 2.2; 20 males). All participants had normal or corrected-to-normal vision and no history of neurological disorders. Prior to the experiment, participants were informed about the experimental procedures, signed informed consent forms, and were given practice trials. Participants received course credit for their participation. Twenty-seven participants performed Experiment A, with mean age 20.4 years (SD = 3.0; 6 males). Forty-six participants performed Experiment B, with mean age 19.8 years (SD = 1.4; 14 males).

2.2. Stimuli

The experiment comprised 408 sentences, which were used in both experiments. All experimental sentences were of the form "Noun-Phrase Verb Noun-Phrase" and either emphasized a first-person or third-person perspective using pronouns *You* or *He/She* (e.g., *You received the letter/He received the letter*). Two hundred and seventy-two experimental sentences were presented (34 sentences per condition, every sentence was presented in the You/He-form). In addition, we presented 68 (Filler) sentences that did not denote any transfer (e.g., *You analyzed the data/He analyzed the data*) and 68 Pseudoword sentences (e.g., *You feaped the mirror/He feaped the mirror*). Sentences belonged to one of four experimental conditions: (1) sentences denoting a concrete transfer towards oneself (e.g., *You plucked the instrument/He plucked the instrument*), (2) sentences denoting a concrete transfer away from oneself (e.g., *You donated the instrument/He donated the instrument*), (3) sentences denoting an abstract transfer towards oneself (e.g., *You heard the instrument/He heard the instrument*), (4) sentences denoting an abstract transfer away from oneself (e.g., *You taught the instrument/He taught the instrument*).

The verbs in Concrete Towards/Away conditions, and similarly the verbs in Abstract Towards/Away conditions, were matched on a number of psycholinguistic variables (see Table 1). They were matched for word frequency per million (Baayen, Piepenbrock, & Van Rijn, 1995) (both ps > .70), word length (both ps > .40), number of phonemes (both ps > .40), and number of syllables (both ps > .60). In addition, using measures from the English Lexicon Project database (Balota et al., 2007), they were also matched for RTs, Familiarity, and Imageability. Furthermore, 26 participants, who did not participate in the main experiment, rated how strongly the verb was associated with an action on a scale from 1 (not associated with action) to 7 (strongly associated with an action), and action-directionality (1 being very strongly related to an action towards the body; 7 being very strongly related to an action away from the body). No difference was observed in the action ratings between Concrete Towards and Concrete Away verbs, $p > .1$. Likewise, no difference was observed in the action ratings between Abstract Towards and Abstract Away verbs, $p > .1$. As expected, Concrete verbs were significantly more action-related than Abstract verbs, $p < .001$. In addition, results of our questionnaire showed that Concrete Towards verbs were associated with a towards action, whereas Concrete Away verbs were associated with an away action, $p < .001$. Likewise, Abstract Towards and Away verbs showed a difference in directionality ($p < .001$).

2.3. Procedure

Participants were seated in front of a computer monitor and responded by means of a key press. In Experiment A, three response buttons were positioned such that a red button was in the middle and blue and yellow buttons were in front of/behind it (see Fig. 1). The buttons were circular with 3” of diameter, and the centers of the buttons were approximately 5” apart, which was held constant across all participants. Participants

Table 1
The psycholinguistic variables for all experimental conditions

	Concrete Towards	Concrete Away	Abstract Towards	Abstract Away
Verb length	5.38 (1.76)	5.29 (1.66)	6.03 (1.55)	6.35 (1.92)
No. of phonemes	4.41 (1.73)	4.29 (1.87)	5.06 (1.65)	5.38 (2.03)
No. of syllables	1.53 (0.66)	1.44 (0.75)	1.94 (0.69)	1.88 (0.69)
RT	654 (68)	628 (56)	654 (64)	679 (89)
Frequency	124.6 (327.5)	113.0 (158.3)	103.6 (138.9)	90.9 (202.9)
Familiarity	6.93 (0.17)	6.92 (0.18)	6.91 (0.16)	6.87 (0.23)
Imageability	4.22 (0.56)	4.34 (0.85)	3.62 (0.91)	3.72 (0.66)
Action rating	5.39 (1.06)	5.75 (0.93)	3.35 (1.02)	3.75 (1.08)
Towards/away	2.24 (1.24)	5.49 (1.09)	2.35 (0.92)	5.44 (0.52)

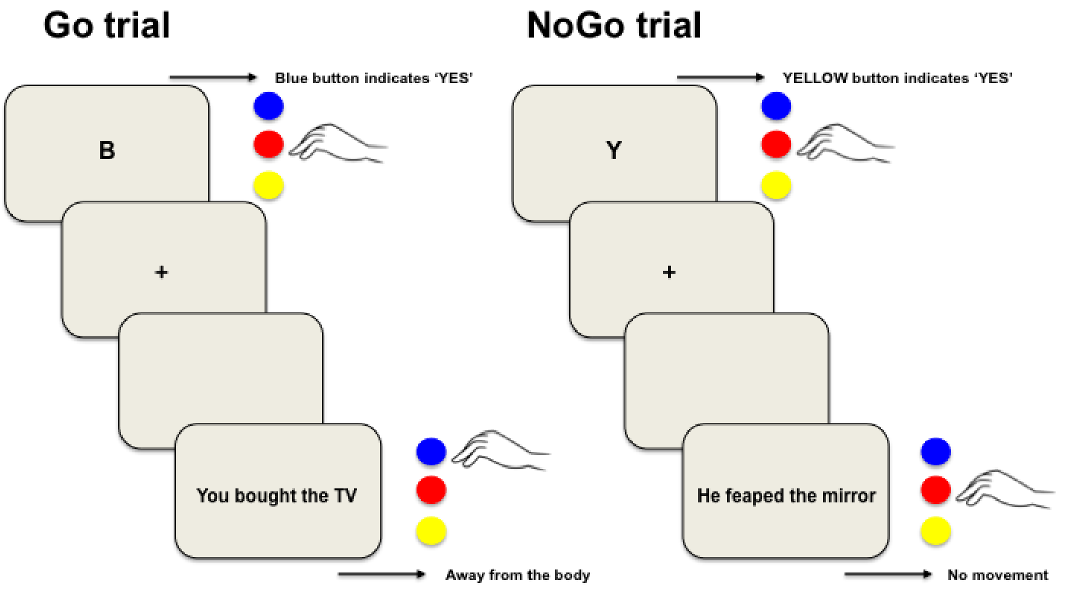

Fig. 1. Illustration of a Go and NoGo trial of Experiment A.

responded by pressing a key that was located *nearer* or *further* from the body. Congruency is labeled in terms of compatibility in terms of movement direction (along front-to-back axis) and implied sentence direction. In Experiment B, the response device was turned 90° such that the red button was again in the middle and blue and yellow buttons were to the left/right of it. The positions of the blue and yellow buttons were counterbalanced across participants for both experiments. Participants responded by pressing a key that was located *left* or *right* from the body. Consistent with previous studies (Chatterjee et al., 1999; Maas & Russo, 2003), congruency was labeled assuming that the agent occupies the left and the patient the right side of conceptual space.

To start a trial, participants had to press the middle button (red button). After 1,500 ms, they received a cue (i.e., either the letter *B* or *Y*) that signaled them to prepare a movement (either to the *Blue* or *Yellow* button), which they only executed if the sentence made sense. This cue was visible for 750 ms; subsequently a fixation cross (appearing centrally for 750 ms) signaled that the sentence would be presented shortly. After a blank screen that was presented for 500 ms, the sentence appeared for 3 s. Sentences were presented in white Arial fonts on a black background and were randomly ordered separately for each subject. Half of the sentences were randomly selected (for each subject separately) to be presented within the congruent versus incongruent condition.

2.4. Data analysis

We measured the latencies to recognize that a sentence was sensible, defined as the time difference between sentence onset and release of the start button. In addition, we recorded movement times (MT: i.e., the time from releasing the start button until depressing the target button). Trials with RTs or MTs (2 * STD ± mean) were treated as outliers and excluded from further analysis. The significance criterion for all analyses was set to $\alpha = 0.05$.

3. Results

The error rates in the sensibility judgment task were on average lower than 2% and therefore not further analyzed. RTs were averaged for each participant in each condition (see Tables 2 and 3; Fig. 2 for means). Reaction times for concrete and abstract sentences were submitted to a separate three-by-two repeated-measures analysis of variance (ANOVA) with the factors of Action Congruency (congruent vs. incongruent), Perspective (first-person perspective vs. third-person perspective), and Response direction (front-to-back vs. left-to-right axis).

For Concrete sentences we observed a three-way interaction (Experiment × Perspective × Congruency), $F(1, 71) = 4.58$, $p < .02$. To further explore this interaction, we examined each experiment separately using post hoc one-sided paired sample *t*-tests. For concrete sentences in Experiment A, congruent sentences were processed faster than incongruent sentences when in first-person perspective ($t(26) = 1.78$, $p < .05$). For the concrete sentences in the third-person perspective, mean reaction times to action congruent and incongruent sentences did not differ statistically, $t(26) = 0.24$, $p > .50$). Thus, a

Table 2
Average reaction times (RTs) with standard errors (*SE*s) of congruent and incongruent trials for concrete sentences, both for the first-person and third-person perspective in Experiments A and B

Concrete Sentences	First-Person Perspective		Third-Person Perspective	
	Front-to-back (Exp. A)	Left-to-right (Exp. B)	Front-to-back (Exp. A)	Left-to-right (Exp. B)
Congruent	1,300 (45)	1,241 (41)	1,304 (43)	1,211 (36)
Incongruent	1,321 (43)	1,229 (38)	1,300 (45)	1,235 (41)

Table 3
Average reaction times (RTs) with standard errors (*SE*s) of congruent and incongruent trials for abstract sentences, both for the first-person and third-person perspective in Experiment A and B

Abstract Sentences	First-Person Perspective		Third-Person Perspective	
	Front-to-Back (Exp A)	Left-to-Right (Exp B)	Front-to-Back (Exp A)	Left-to-Right (Exp B)
Congruent	1,372 (52)	1,276 (40)	1,371 (51)	1,267 (43)
Incongruent	1,361 (49)	1,280 (40)	1,373 (47)	1,270 (39)

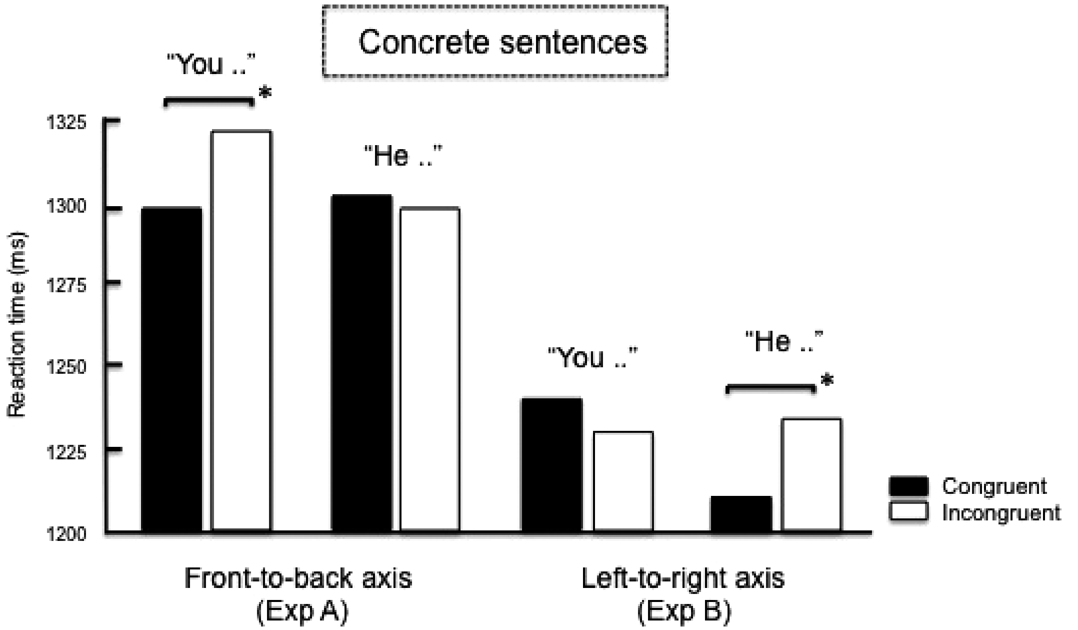

Fig. 2. Average reaction times (RTs) for sentences, as a function of the congruency between the cue and the associated movement direction and the orientation of the response device for concrete sentences in the first-person and third-person perspective.

reliable ACE was obtained for sentences that described a concrete transfer from the first-person perspective, whereas no such effect was observed for sentences that implied a third-person perspective. These findings suggest that the form of utterances can affect the perspective that readers adopt in the motoric information they activate in their mental simulations.

It could be the case that for sentences that imply a third-person perspective (e.g., "*He threw the microphone*"), readers take an external perspective and simulate those motor actions from an observer's position. If this is the case, it can be expected that a congruency effect can be observed with *left/right* instead of *towards/away* motor responses. Likely, the actor's position in the linguistic utterance will guide the mental simulation (i.e., activation of a [left to] *right* motor code for the sentence *He threw the microphone*). We found that in Experiment B, for concrete sentences in the third-person perspective, congruent sentences were processed faster than incongruent sentences, $t(26) = 1.82$, $p < .05$. For the concrete sentences in the first-person perspective, mean reaction times to action congruent and incongruent sentences did not differ statistically, $t(26) = 0.91$, $p > .80$. Thus, an ACE was obtained for sentences that described a concrete transfer from the third-person perspective, whereas no such effect was observed for sentences that implied a first-person perspective.

For Abstract sentences the three-way interaction (Experiment × Perspective × Congruency) failed to reach significance, $F(1, 71) = 0.15$, $p > .35$. No differences were observed between the action congruent and incongruent Abstract sentences in the First-Person or Third-Person Perspective of Experiment A and B (all ps > .35).

4. Discussion

The current experiment explored if facilitation of sentence comprehension by prior planning of an action congruent to the movement direction implied in the sentence is modulated by sentence perspective. Previous studies have shown that the form of an utterance and the perspective imposed on objects by a linguistic utterance modulate the perspective that comprehenders adopt in their mental simulation of perceptual attributes of objects and events (Borghi et al., 2004; Brunyé et al., 2009; Ditman et al., 2010). This study shows that the motor attributes that comprehenders activate in their mental simulations of described events are similarly modulated by the form of linguistic phrases. Sentences that described a transfer of an object from the first-person perspective showed an ACE along the *towards-away* axis, whereas sentences from the third-person perspective demonstrated an ACE along the left-right axis. That is, the same verb "*throw*" triggered a *towards-away* simulation when paired with a first-person agent ("*You threw the microphone*"), whereas it triggered a left-right simulation when paired with a third-person agent ("*He threw the microphone*"). These results corroborate findings by Brunyé et al. (2009) that showed that differences in the form of utterances affect the perspectives that comprehenders adopt in their mental simulation of perceptual attributes of objects and events.

Finding an ACE along the left-right axis for sentences from the third-person perspective substantiates the idea that conceptual ordering of agents and patients to the left and right of space, respectively, at least for English, matches its orthographic direction (Chatterjee et al., 1999; Maas & Russo, 2003). If agents and patients in real life do not correspond to any side of space, then why should there be a bias to put agents on the left and patients (objects) on the right? Chatterjee, Maher, Rothi, and Heilman (1995) described a left-handed patient, WH, who suffered a right hemisphere infarct causing agrammatic aphasia. For both active and passive reversible sentences, WH was unable to map agents/patients onto grammatical subjects/objects. He, however, performed at ceiling when actions were depicted going from left to right for active sentences, and from right to left for passive sentences (see also Boiteau & Almor, 2016). A similar spatiotemporal bias in thematic role assignment has been observed in healthy subjects. Chatterjee et al. (1999) presented sentences for which subjects had to draw the agent and patient. They showed that subjects consistently drew agents to the left and patients to the right of space. Subjects were also faster at matching sentences to pictures if the picture depicted the agent of a transitive action on the left and patient on the right. Maas and Russo (2003) investigated whether this bias is due to orthographic direction, comparing Italian (left-to-right orthography) and Arabic (right-to-left) subjects using a sentence-picture matching task. Participants heard sentences that either implied a subject-to-object (e.g., "*The boy/girl pushes the girl/boy*") or object-to-subject motion (e.g., "*The boy/girl pulls the girl/boy*"). Images displayed the subject either on the left or the right side of the screen and were paired with an action that either had a left-to-right flow (subject on the left pushing object; subject on the right pulling object) or right-to-left flow. For the trials in which the sentence and picture matched, Italian speakers were facilitated by images showing agents on the left, whereas Arabic speakers by agents on the right, suggesting that orthographic direction is indeed a driver of this effect. However, both groups showed a faster response to images showing a left-to-right flow, suggesting that language laterality likely also contributes to the effect. Chatterjee (2001) proposed that most individuals (for whom the left hemisphere is language dominant) process transitive actions via the left temporo-parietal-occipital junction, which plays a role in left-to-right motion tracking. Left hemisphere regions are also involved in launching contraversive saccades, which facilitate left-to-right scanning and linguistic processing (Herter & Guitton, 2004). On the basis of this work on spatial biases in thematic role assignment, we defined the compatibility between sentences and responses in our study (i.e., agents taking the left side of conceptual space and the direction of motion specified by the main verb). That is, our finding of facilitated left-to-right responses for sentences with a main verb like "*throw*" and right-to-left responses for sentences with a main verb like "*catch,*" suggests that we defined compatibility correctly (although a facilitation in the opposite direction would also be sufficient to demonstrate that third-person sentence comprehension involves visuo-motor simulations).

In the original work on spatial alignment effects, an overlap of low-level visual features of the stimulus material and response characteristics has been pointed out as a possible confound (Matheson, White, & McMullen, 2014). Furthermore, spatial alignment effects have been demonstrated for objects that do not typically afford actions (e.g., the hands of a clock) (Anderson, Yamagishi, & Karavia, 2002). Spatial alignment effects have also been shown when responses were made with different fingers from the same hand or the feet (Cho & Proctor, 2010; Phillips & Ward, 2002). On the basis of these findings, several authors have questioned the sensory-motor nature of spatial alignment effects and have argued for an explanation in terms of a general stimulus-response compatibility, or as an instance of the Simon effect. However, a number of studies have found similar compatibility effects for objects that afford different types of grasps rather than effector-specific object affordances (Bub et al., 2008; Helbig et al., 2010; Masson et al., 2011; Tucker & Ellis, 2001). These grasp compatibility effects have also been obtained if the names of the objects were presented (Bub et al., 2008; Tucker & Ellis, 2004), suggesting a conceptual nature of the effect regardless of the mode of access (i.e., visual vs. linguistic). These findings cannot easily be explained in terms of a general stimulus-response compatibility and point towards an involvement of sensory-motor processes in conceptual representations accessed through pictures. It should be noted that evidence for sensory-motor activations for objects that afford different types of grasps does not prove a similar sensory-motor nature of spatial alignment effects, which may result from general stimulus-response compatibility. Importantly, this study is fundamentally different from the aforementioned studies in that the thematic roles and the position of the subject/object were kept constant across conditions. The effects arise due to changes in verb semantics, not due to changes in the physical position of the stimulus.

Thus, these effects cannot be interpreted in terms of the Simon effect. Furthermore, if compatibility effects would be purely driven by the position of patients/agents or subjects/objects within a sentence, then we would expect a similar compatibility effect for first-person sentences if making left-right responses, which is not the case.

The results of this study suggest that, similar to first-person sentences, participants simulate the movement direction implied in the main verb for third-person sentences. A recent study by Secora and Emmorey (2015) examined the nature of sensory-motor simulations by studying the ACE in deaf American Sign Language (ASL) signers. In sign languages, the semantics of a verb like "throw" implies a movement away from the body, whereas the sign for the main verb involves a visual movement towards the addressee. Deaf ASL signers were asked to perform a sentence sensibility judgment, whereas signed sentences were addressed to them that expressed either toward (e.g., "*You grabbed a cup*") or away motion (e.g., "*You threw a ball*"). They found that the direction of motion of the signer's hand with respect to the addressee's body did not modulate the ACE (i.e., no congruency effect was observed when responses were categorized in relation to the perceptual motion of the sentence). These results suggest a motoric nature of the mental simulation during the processing of a first-person sentence like "*you threw a ball.*" Furthermore, neuroimaging studies have demonstrated an involvement of primary and higher order motor areas in the processing of similar literal action sentences (Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006; Boulenger, Hauk, & Pulvermüller, 2009; Desai, Binder, Conant, & Seidenberg, 2010; Van Dam & Desai, 2016). The interaction of our compatibility effect with response direction (front-to-back vs. left-to-right axis) shows that perspective information is incorporated in these mental simulations. However, it can be argued that in the case of third-person sentences the nature in which the representational content is grounded might be more strongly based on visual representations. That is, a perceptual simulation of a movement of an object from one location to another rather than a simulation of the motoric actions that causes these movements.

A study by Gianelli, Farnè, Salemme, Jeannerod, and Roy (2011) investigated whether the ACE typically observed for a first-person sentence like "*You gave x to Louis*" would extend to third-person sentences like "*Lea gave x to Louis*" if readers are instructed to take the perspective of the agent. They observed an ACE, but only if subjects were given an anchor that provided them with information "*where*" to position themselves in space when taking the perspective of the agent. An ACE was observed along the front-back rather than left-right axis for third-person sentences, if readers took the perspective of the actor in the sentence. One could speculate that taking the perspective of the agent in these third-person sentences, effectively converting them to first-person sentences, may have activated front-back effects that compete with the default right-left effects induced by the third-person sentences. In addition, representations evoked by the task are likely imagery-based, given that their instruction *to perform the task as if* they were the agent encouraged subjects to perform mental imagery of the linguistically denoted actions. Such imagery-based representations might only be strong enough to overwrite the spatial bias that arises due to simulating the perspective that is inherently implied in the sentence, if subjects are provided with an additional spatial anchor to guide their perspective taking. Thus, their results might reflect a competition between simulations of the perspective that is inherently implied in the sentence and task-induced representations. This study differs in terms of the stimulus materials (motion direction implied by the verb rather than by the sentence structure) and task (no direct or indirect instructions to take a perspective or perform imagery).

We did not observe an ACE for abstract sentences in either experiment. This finding might reflect the fact that words and sentences with abstract meanings do not tend to activate the neural motor system (Cacciari & Pesciarelli, 2013; Cacciari et al., 2011; Raposo et al., 2009; Rueschemeyer et al., 2007) and therefore are immune to showing an ACE. Glenberg and Kaschak (2002) documented an ACE for sentences describing an abstract transfer. The difference with our study may be due to differences in the linguistic material used in both studies. Half of the sentences in the Glenberg and Kaschak (2002) study used the dative form (e.g., *The policeman radioed the message to you*) and the other half of the sentences used the double object construction (e.g., *Liz told you the story*). These more elaborate sentence structures, which are characterized by a verb, an agent argument, a recipient-like argument, and a theme argument, likely put a stronger emphasis on directional transfer and therefore even show effects for syntactic constructions describing abstract transfer. That is, in these examples the syntactic constructions themselves might elicit sensory-motor activations, independent of the main verb used in the constructions. Several studies have provided evidence that syntactic constructions themselves can carry meaning (Fisher, 1996; Kako, 2006; Kaschak & Glenberg, 2000; Van Dam & Desai, 2016). Therefore, comprehenders might activate perceptual and motor information in deriving lexical-semantic meaning at the level of syntactic constructions (Allen, Pereira, Botvinick, & Glenberg, 2012). In the Glenberg and Kaschak (2002) study, the dative/double object constructions likely emphasized movement direction resulting in an ACE. In our study, on the other hand, information about movement direction was provided by the main verbs in the sentences which were abstract in nature, and which have shown to differ in their representational content from concrete concepts (Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011; Wiemer-Hastings & Xu, 2005). That is, sensory-motor activation at the level of syntactic constructions in these simple transitives (NP vs. NP) might be weaker. In the case that movement direction is denoted by an abstract main verb, this might invoke a more abstracted simulation that contains weaker directional and perspective information.

5. Conclusion

We demonstrated that the pronominal perspective implied in the sentence modulates the interaction between sentence comprehension and movement preparation processes. Our findings provide evidence that in sentence comprehension, perspective is part of the mental simulations of motor attributes of described events.

Acknowledgments

This research was supported by NIH/NIDCD grant R01 DC010783. We thank Nick Riccardi for his assistance in running subjects.

References

Allen, K., Pereira, F., Botvinick, M., & Glenberg, A. E. (2012). Distinguishing grammatical constructions with fMRI pattern analysis. *Brain and Language*, *123*(3), 174–182.

Anderson, S. J., Yamagishi, N., & Karavia, V. (2002). Attentional processes link perception and action. *Proceedings of the Royal Society of London*, *269*(1497), 1225–1232.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, *16*, 1818–1823.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1995). The CELEX Lexical Database. Release 2 [CD-ROM]. Linguistic Data Consortium, University of Pennsylvania, Philadelphia.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*, 445–459.

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

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Pyschology*, *59*, 617–645.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*(11), 527–536.

Black, J. B., Turner, T. J., & Bower, G. H. (1979). Point of view in narrative comprehension, memory, and production. *Journal of Verbal Learning and Behaviour*, *18*, 187–198.

Boiteau, T., & Almor, A. (2016). Transitivity, space, and hand: The spatial grounding of syntax. *Cognitive Science*, 1–44, doi: 10.1111/cogs.12355

Borghi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition*, *32*(6), 863–873.

Boulenger, V., Hauk, O., & Pulvermüller, F. (2009). Grasping ideas with the motor system: Semantic somatotopy in idiom comprehension. *Cerebral Cortex*, *19*(8), 1905–1914.

Brunyé, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. A. (2009). When you and I share perspectives: Pronouns modulate perspective taking during narrative comprehension. *Psychological Science*, *20*(1), 27–32.

Bub, D. N., Masson, M. E. J., & Cree, G. S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition*, *106*, 27–58.

Cacciari, C., Pellicciari, M. C., Bolognini, N., Senna, I., Miniussi, C., & Papagno, C. (2011). Literal, fictive and metaphorical motion sentences preserve the motion component of the verb. A TMS study. *Brain and Language*, *119*, 149–157.

Cacciari, C., & Pesciarelli, F. (2013). Motor activation in literal and non-literal sentences: Does time matter? *Frontiers in Human Neuroscience*, *7*(202), 1–10. doi:10.3389/fnhum.2013.00202

Chatterjee, A. (2001). Language and space: Some interactions. *Trends in Cognitive Science*, *5*(2), 55–61.

Chatterjee, A., Maher, L. M., Rothi, L. J. G., & Heilman, K. M. (1995). Asyntactic thematic role assignment: The use of a temporal-spatial strategy. *Brain and Language*, *49*(2), 125–139.

Chatterjee, A., Southwood, M. H., & Basilico, D. (1999). Verbs, events, and spatial representations. *Neuropsychologia*, *37*, 395–402.

Cho, D., & Proctor, R. W. (2010). The object-based Simon-effect: Grasping affordance or relative location of the graspable part? *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 737–749.

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

Desai, R. H., Binder, J. R., Conant, L. L., & Seidenberg, M. S. (2010). Activation of sensory-motor areas in sentence comprehension. *Cerebral Cortex*, *20*(2), 468–478.

Desai, R. H., Conant, L. L., Binder, J. R., Park, H., & Seidenberg, M. S. (2013). A piece of the action: Modulation of sensory-motor regions by action idioms and metaphors. *NeuroImage*, *83*, 862–869.

Ditman, T., Brunyé, T. T., Mahoney, C. R., & Taylor, H. A. (2010). Simulating an enactment effect: Pronouns guide action simulation during narrative comprehension. *Cognition*, *115*, 172–178.

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology*, *23*, 331–392.

Fischer, M., & Zwaan, R. (2008). Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology*, *61*, 825–850.

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Gianelli, C., Farnè, A., Salemme, R., Jeannerod, M., & Roy, A. C. (2011). The agent is right: When motor embodied cognition is space-dependent. *PLoS ONE*, *6*(9), 1–8.

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

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

Helbig, H., Steinwender, J., Graf, M., & Kiefer, M. (2010). Action observation can prime visual object recognition. *Experimental Brain Research*, *55*, 270–279.

Herter, T. M., & Guitton, D. (2004). Accurate bidirectional saccade control by a single hemicortex. *Brain*, *127*, 1393–1402.

Kako, E. (2006). The semantics of syntactic frames. *Language and Cognitive Processes*, *21*(5), 562–575.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory & Language*, *43*, 508–529.

Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language*, *39*, 170–210.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, *48*, 805–825.

Kintsch, W. (2008). Symbol systems and perceptual representations. In M. De Vega, A. Glenberg, & A. Graesser (Eds.), *Symbols and embodiment* (pp. 14–164). Oxford, UK: Oxford University Press.

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, *140*(1), 14–34.

Maas, A., & Russo, A. (2003). Directional bias in the mental representation of spatial events: Nature or culture? *Psychological Science*, *14*(4), 296–301.

Matheson, H. E., White, N. C., & McMullen, P. A. (2014). A test of the embodied simulation theory of object perception: potentiation of responses to artifacts and animals. *Psychological Research*, *78*(4), 465–482.

Masson, M. E. J., Bub, D. N., & Breuer, A. T. (2011). Priming of reach and grasp actions by handled objects. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1470–1484.

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, *48*(7), 788–804.

Morrow, D. G., & Clark, H. H. (1988). Interpreting words in spatial descriptions. *Language and Cognitive Processes*, *3*, 275–291.

Phillips, J. C., & Ward, R. (2002). S-R correspondence effects of irrelevant visual affordance: Time course and specificity of response activation. *Visual Cognition*, *9*, 540–558.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral Brain Sciences*, *22*(2), 253–336.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*, 576–582.

Pustejowsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.

Pylyshyn, Z. W. (1984). *Computation and cognition: Towards a foundation for cognitive science*. Cambridge: MIT Press.

Raposo, A., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia*, *47*, 388–396.

Rueschemeyer, S.-A., Brass, M., & Friederici, A. D. (2007). Comprehending prehending: Neural correlates of processing verbs with motor stems. *Journal of Cognitive Neuroscience*, *19*, 855–865.

Rueschemeyer, S.-A., Lindemann, O., Van Rooij, D., Van Dam, W. O., & Bekkering, H. (2010). Effects of intentional motor actions on embodied language processing. *Experimental Psychology*, *57*(4), 260–266.

Sanford, A. J., & Garrod, S. C. (1998). The role of scenario mapping in text comprehension. *Discourse Processes*, *26*(2&3), 159–190. doi:10.1080/01638539809545043.

Schwarzkopf, S., Weldle, H., Müller, D., & Konieczny, L. (2011). Mental simulation of spatial perspective during sentence comprehension. Paper presented at the 33rd Annual Conference of the Cognitive Science Society (CogSci 11), July 2011, Boston.

Secora, K., & Emmorey, K. (2015). The action-sentence compatibility effect in ASL: The role of semantics vs. perception. *Language and Cognition*, *7* (2), 305–318.

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

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 830–846.

Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta Psychologica*, *116*(2), 185–203.

Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition*, *8*(6), 769–800.

Van Dam, W. O., & Desai, R. H. (2016). The semantics of syntax: The grounding of transitive and intransitive constructions. *Journal of Cognitive Neuroscience*, *28*(5), 693–709.

Van Dam, W. O., Rueschemeyer, S.-A., Lindemann, O., & Bekkering, H. (2010). Context effects in embodied lexical-semantic processing. *Frontiers in Psychology*, *1*, 150. doi:10.3389/fpsyg.2010.00150.

Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive Science: A Multidisciplinary Journal*, *29*, 719–736.

Zwaan, R. A., & Kaschak, M. P. (2008). Language in the brain, body, and world. In M. Robbins & M. Aydede (Eds.), *Cambridge handbook of situated cognition* (pp. 368–381). Cambridge, UK: Cambridge University press.

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