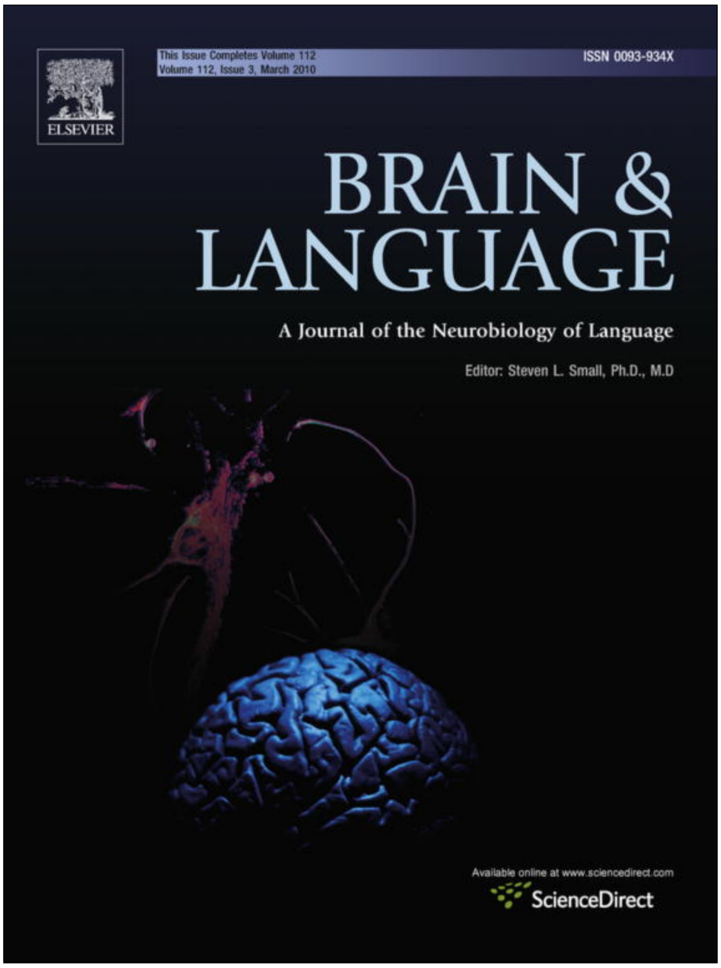

Brain & Language 112 (2010) 150–158

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Grammatical aspect and mental simulation

Benjamin Bergen *, Kathryn Wheeler

University of Hawai'i at Manoa, Department of Linguistics, 569 Moore Hall, 1890 East-West Rd., Honolulu, HI 96822, United States

ARTICLE INFO

Article history:
Accepted 3 July 2009
Available online 4 August 2009

Keywords:
Mental simulation
Motor simulation
Mental imagery
Language comprehension
Cognitive linguistics
Grammatical aspect

ABSTRACT

When processing sentences about perceptible scenes and performable actions, language understanders activate perceptual and motor systems to perform mental simulations of those events. But little is known about exactly what linguistic elements activate modality-specific systems during language processing. While it is known that content words, like nouns and verbs, influence the content of a mental simulation, the role of grammar is less well understood. We investigate the role of grammatical markers in mental simulation through two experiments in which we manipulate the meanings of sentences by modifying the grammatical aspect they use. Using the Action-sentence Compatibility Effect (ACE) methodology [Glenberg, A., Kaschak, M. (2002). Grounding language in action. *Psychonomic Bulletin and Review, 9*, 558–565], we show that progressive sentences about hand motion facilitate manual action in the same direction, while perfect sentences that are identical in every way except their aspect do not. The broader implication of this finding for language processing is that while content words tell understanders what to mentally simulate and what brain regions to use in performing these simulations, grammatical constructions such as aspect modulate how those simulations are performed.

1. Introduction

Processing language uses perceptual and motor areas of the brain, not only to perceive written characters or spoken sounds, but also to extract meaning from utterances. When people process language that describes perceivable scenes or performable actions, they display selective activation of perceptual and motor systems (Pulvermüller, Haerle, & Hummel, 2001; Hauk, Johnsrude, & Pulvermuller, 2004). For instance, processing sentences about motor actions using the hand, like *You are giving Andy a slice of pizza*, activates a portion of the motor strip dedicated to hand motion, while sentences that describe foot actions like *You are kicking the ball across the field* lead to activation of areas responsible for controlling the foot (Tettamanti et al., 2005; Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006).

In the literature on language understanding, this modality-specific activation (that is, activation of brain systems dedicated to perception or action) has been interpreted as reflecting *mental simulation* of the content of the sentences on the part of understanders (Glenberg & Kaschak, 2002; Zwaan, Stanfield, & Yaxley, 2002; Richardson, Spivey, Barsalou, & McRae, 2003; Bergen, Lindsay, Matlock, & Narayanan, 2007). Mental simulation is the internal enactment or reenactment of perceptual, motor, or affective experiences (Barsalou, 1999). Mental simulation may be static or dynamic, and this term is often used interchangeably in the literature with *mental imagery*. Mental simulation is produced by brain structures specific to the relevant modality; motor simulation uses motor areas, down to the specific regions that control simulated effectors (Porro et al., 1996; Lotze et al., 1999; Ehrsson, Geyer, & Naito, 2003; see also the overview in Willems & Hagoort, 2007). Analogously, visual simulation is produced through activation of visual areas (for a review of the evidence, see Kosslyn, Ganis, & Thompson, 2001). Importantly, while the motor and perceptual content of mental simulations can often be made accessible to conscious introspection, mental simulation constructed during language processing is immediate, automatic, and almost entirely unconscious (Barsalou, 1999).

There is ample evidence that processing sentences leads people to perform mental simulation. Hearing a sentence like *You are giving Andy a slice of pizza* can lead understanders to perform a motor simulation of what it would be like to move their hand in an appropriate way to perform the described action (Glenberg & Kaschak, 2002), to form their hand into the appropriate shape (Wheeler & Bergen, in press) or to perform a visual simulation of which direction such a transfer would move in (Richardson et al., 2003; Kaschak et al., 2005). A number of authors have taken these observations further and argued that the use of motor or perceptual systems to perform mental simulation is an integral part of understanding language (Glenberg & Kaschak, 2002; Bergen & Chang, 2005; Feldman & Narayanan, 2004; Gallese & Lakoff, 2005).

But while the existing research on mental simulation during language processing clearly shows that activation of language

---

* Corresponding author. Fax: +1 808 956 9166.
E-mail address: bergen@hawaii.edu (B. Bergen).

0093-934X/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandl.2009.07.002

centers produces activation in perceptual and motor areas (Pulvermuller, 2003), there is little empirical evidence on exactly what components of an utterance drive what aspects of this perceptual and motor simulation. Is it just words, or do grammatical structures contribute as well, and if so, in what ways?

To begin with words, it is well known that as individual lexical representations become active during language processing, activation spreads from them to perceptual and motor representations of the modality-specific experiences that they refer to. This entails associations between representations for words and representations for the percepts and actions associated with whatever they denote. These associations might arise through straightforward Hebbian learning (Colunga & Smith, 2005). For instance, a child might hear the word *hit* in close enough temporal proximity to the experience of actual hitting (motorically, visually, or perhaps haptically) to reinforce connections between their internal representation of the word and of the perceptual or motor experience (Chang, 2004). Later, that same child hears the word *hit*, which directly triggers activation of those same perceptual and motor structures, through strengthened associative connections.

Lexical associations, where activation of word meanings drives perceptual and motor simulation, are consistent with most evidence for mental simulation during language processing. For instance, in their work on spatial simulation, Richardson et al. (2003) took directionally associated words (for instance, *push* tends to denote horizontal motion while *bomb* usually denotes vertical motion), placed these verbs in sentences, and measured the extent of location-specific interference on visual processing. They found that vertically associated verbs slowed visual processing of objects in the vertical axis of a computer screen, while horizontally associated verbs slowed visual processing of objects in the horizontal axis of the screen. This location-specific interference of language on vision is interpreted to be the product of language-induced mental simulation, which ties up location-specific parts of the vision system. This simulation appears to be driven by associations between the words in the sentence and perceptual and motor aspects of their denotations.

But there is evidence that lexical associations are not the whole story. Subsequent to Richardson et al.'s study, other behavioral work focusing on visual processing during language understanding (Bergen et al., 2007) has found that concrete words describing upwards or downwards motion (like *climb* or *fall*) only sometimes produce location-specific processing by the vision system. When those concrete words are embedded in sentences describing actual motion (such as *The mule climbed*), they indeed show reliable activation of the vision system, but not when they are placed in sentences describing metaphorical motion (such as *The prices climbed*). In other words, the word *climb* sometimes does and sometimes does not produce visual processing of imagined upwards motion. A very similar result is reported by Sato, Mengarelli, Riggio, Gallese, and Buccino (2008), who found that different participant instructions (to perform a semantic judgment or a grammaticality judgment) affected whether or not hand and foot action words primed manual action. Brain imaging work also suggests that there is more at work than lexical associations. Aziz-Zadeh et al. (2006) report on a similar difference between literal and metaphorical uses of motor language. Listeners showed somatotopic activation of premotor cortex when exposed to sentences describing real actions involving the foot, hand, or mouth, but not when exposed to metaphorical sentences describing abstract events using foot, hand, or mouth actions verbs. In sum, the larger context, especially the sentential context, seems to be a key factor in determining what perceptual or motor content is mentally simulated.

This observation has led to a more nuanced account of how language processing leads to motor and perceptual activation. Instead of mental simulation being driven just by connections from lexical representations to perceptual and motor areas, a number of authors have theorized a prominent role for grammar (Kaschak & Glenberg, 2000; Bergen & Chang, 2005; Feldman, 2006). Grammar clearly has a role in assembling the contributions of a sentence's lexical parts. But some approaches to language also suggest (Langacker, 1987; Lakoff, 1987; Goldberg, 1995; Talmy, 2000), and experimental studies have confirmed (Kaschak & Glenberg, 2000; Bencini & Goldberg, 2000) that, like words, grammar can affect the content of a mental simulation on its own right.

Consider sentences that differ only in their grammatical markings, like (1a) versus (1b). These particular sentences use the same content words, but different *grammatical aspect* (Comrie, 1976; Dowty, 1977). The use of *progressive aspect* (1a) has been argued to accentuate the internal structure of an event, while *perfect aspect* (1b), has been claimed to encapsulate or shut off access to the described process, while highlighting the resulting endstate (Comrie, 1976; Dowty, 1977; Narayanan, 1997; Chang, Gildea, & Narayanan, 1998). Different methods have confirmed this intuition (Carreiras, Carriedo, Alonso, & Fernández, 1997; Magliano & Schleich, 2000; Madden & Zwaan, 2003; Ferretti, Kutas, & McRae, 2007; Madden & Therriault, 2009). Translated into predictions for mental simulation, we might expect that grammatical aspect would modulate the part of an event that is mentally simulated in the greatest detail. This notion is viable in principle, as the mental simulation literature is rich with demonstrations that mental focus can be placed on different parts of a simulated scene (Denis & Kosslyn, 1999; Mellet et al., 2002; Borghi, Glenberg, & Kaschak, 2004).

(1) a. John is closing the drawer.
b. John has closed the drawer.

Whether grammar significantly affects the content of mental simulations is of prime importance to those theories of sentence processing that view the activation of modality-specific systems in mental simulation as a critical part of understanding (such as Kaschak & Glenberg, 2000; Bergen & Chang, 2005; and Feldman, 2006). The reasoning goes like this. If mental simulation, and the perceptual and motor processing that underlies it, is constitutive of understanding, then this simulation should move in lockstep with the meaning of the utterance being processed. Two sentences with different meanings should produce two measurably different activity patterns in modality-specific systems. This is a critical claim of these approaches to language processing; if sentences with different meanings produce the same activation in modality-specific systems, then mental simulation cannot be integral to – let alone constitutive of – actually understanding language. As a result, if grammatical differences like the progressive/perfect distinction produce no affect on mental simulation – if the words in an utterance activate associated motor and perceptual representations on a word-by-word basis – then activation of motor or perceptual systems is not sufficient for understanding.

To test this claim, we took sentences like those in (1). We hypothesized that progressive aspect, as in (1a), which should focus mental simulation on the central action (the act of pushing on the drawer), will yield measurable activation of motor systems. By contrast, perfect aspect (1b) should lead the understander to focus mental simulation on the endstate of the event – the drawer in its final, closed position – and, as a result, will not activate motor simulation of the closing action. If sentences using the same content words but these different grammatical constructions produce systematically different use of the motor system, then this suggests that the motor system is only activated when the words and the grammar of an utterance conspire to not only indicate a motor action event but also focus on the processes of that motor action.[1]

To test this hypothesis, we conducted an Action-sentence Compatibility Effect experiment (Glenberg & Kaschak, 2002), where participants pressed a button – which was located in the middle of a keyboard – to trigger the visual presentation of a sentence on the screen. When they released the button, the sentence disappeared, and they then pressed a second button to indicate whether the sentence was meaningful or not. Critically, the second button was located either closer to or farther from the participant's body than the first, so pressing it required them to make a hand movement either towards or away from their body. Previous studies (Glenberg & Kaschak, 2002; Bergen & Wheeler, 2005; Tseng & Bergen, 2005; Borreggine & Kaschak, 2006) have shown that when the direction of motion described by a sentence is the same as the direction of the response arm movement, participants perform faster manual responses, and this is indicative that they have used the motor system to mentally simulate manual action in the direction denoted by the sentence.

The key independent variable was whether the direction of the participant's response action was compatible or incompatible with the direction of action described in the sentences. To test for effects of grammatical aspect on mental simulation, we conducted two experiments, which differed only in the aspect of the stimuli. Participants in Experiment 1 read progressive sentences, as in (2). We expected to find a significant Action-sentence Compatibility Effect in this experiment. Participants in Experiment 2 read perfect sentences, as in (3), which were hypothesized not to yield simulation detail pertaining to the actual motor performance of the action.[2]

(2) a. John is closing the drawer.
b. John is opening the drawer.

(3) a. John has closed the drawer.
b. John has opened the drawer.

In this way, the two experiments allowed us to test the prediction that progressive sentences will drive the motor system to perform mental simulation of the core or nucleus of the described action, but perfect sentences will not.[3] This is precisely what we found, showing that grammar plays a role in determining whether and how projections from lexical representations in language centers to modality-specific regions where mental simulation is enacted are used.

2. Experiment 1: progressive

2.1. Participants and materials

Seventy University of Hawai'i at Manoa students participated in exchange for either course credit in an introductory linguistics class or $5. All were right-handed native English speakers.

A total of 200 sentences were created: 80 meaningful critical sentences, 40 meaningful filler sentences, and 80 non-meaningful filler sentences. The 80 critical sentences (in the Appendix) were composed of 40 pairs of sentences. In each pair, one sentence denoted a hand motion forwards, away from the body and the other denoted a hand motion backwards, towards the body. These 80 critical stimuli were of two types. One set of 40 consisted of 20 pairs of transitive sentences that critically differed only in their object noun phrase (4a). The second set consisted of 20 pairs of transitive sentences that critically differed only in their main verb (4b). We expected these two sets of sentences, which both described literal hand actions towards or away from the body, to yield similar Action-sentence Compatibility Effects. However, we separated them out for analysis in order to observe any eventual differences. All referents in all sentences were third-person. In this Progressive experiment, all sentences were in the present progressive tense (4).

(4) a. Richard is beating (the drum/his chest).
b. Carol is (taking off/putting on) her glasses.

Sentence pairs were drawn (with some modifications) from the stimuli used by Glenberg and Kaschak (2002), in addition to newly generated ones conforming to the criteria described above. In all, we generated 25 sentences per condition. These potential stimuli were then submitted to a norming study in order to choose pairs whose members encoded the appropriate direction of motion. In the norming study, 12 participants, all native speakers of English, were instructed to decide if the described action required movement of the hand toward or away from the body. To respond, they pressed one of three buttons, labeled *toward*, *away*, or *neither*. We coded the frequency of each response. In order to select stimuli for the main experiment, we excluded sentences that received less than 50% agreement on a single motion direction, and those that had more than 25% of responses in the opposite direction. Of the remaining stimuli, we selected 20 per condition, seen in the Appendix. The mean agreement on the predicted motion direction for critical stimuli as a whole was 72% (standard deviation = 17%), while mean responses in the opposite direction was 10% (standard deviation = 7%). We were thus confident that the sentences were systematically evoking motion in the intended direction.

2.2. Design and procedure

Each participant saw 160 sentences, composed of all 120 fillers and one sentence from each of the 40 critical pairs. Each run of the experiment was split into two halves. For all participants, the 'YES' button was farther from them and the 'NO' button was closer to them in the first half. The button assignments were switched for the second half of the experiment. For each participant, the direction of critical sentences (toward the body and away from the body) was crossed with Response-Direction (YES-is-far or YES-is-close) by placing half of the critical sentences in each of two halves of the experiment. This produced four versions of the experiment, and each participant was randomly assigned to one of the four versions prior to beginning the experiment. Thus, each sentence was answered in the YES-is-far condition by half of the participants ($n = 35$) and in the YES-is-close condition by the other half ($n = 35$).

For response collection, a standard personal computer keyboard was rotated 90° counter-clockwise so that it lay in front of the participant along their sagittal axis. In each trial, participants first saw a fixation cross, at which point they pressed and held a yellow button (the *h* key in the middle the keyboard) to reveal a written sentence until they had decided if (YES) the sentence made sense or (NO) it did not, whereupon they re-

---

[1] This reasoning does not imply an absence of mental simulation in the processing of perfect sentences. Madden & Zwaan's (2003) study, for instance, makes a clear case for the presence of visual simulation during perfect sentence processing. The claim is simply that progressive sentences drive understanders to simulate the core or nucleus of an event (whether through motor, perceptual, or other simulation), while perfect sentences do not – naturally, there are many other things for people processing perfect sentences to simulate in any of a number of modalities.

[2] We used the present participle *has Xed* because it is an unambiguous marker of perfective aspect, unlike the simple past *Xed*, and also because it is matched for tense (present) with the present progressive *is Xing*.

[3] We manipulated aspect between two experiments because there is a limited set of viable stimuli describing actions that systematically evoke hand motion towards or away from the body. A between participants design allowed us to avoid carryover effects resulting from multiple presentations of similar sentences to a single participant, or insufficient power caused by low numbers of stimuli in each condition.

leased the yellow button and pressed a button labeled 'YES' or 'NO' (the *a* or ' key). Participants were instructed to use only their right hand during the experiment. Because the key assignments changed between the two blocks, a training session of 10 trials preceded each half of the experiment.

There are three measures of participants' responses that have shown Action-sentence Compatibility Effects. The first, reported by Glenberg and Kaschak (2002), is on the time it takes participants to read the sentence and then release the middle button. The second is on the time it takes participants to subsequently press the proximal or distal YES button to indicate that the sentence is meaningful (Bergen & Wheeler, 2005; Tseng & Bergen, 2005). Third, the effect can appear on the aggregate of these two (Borreggine & Kaschak, 2006). One factor that seems to influence where the effect is observed is whether sentences include the word *you* or not. In studies in which sentences describe actions either performed by *you* or on *you*, the effect appears on the earlier measure of middle-button release (Glenberg & Kaschak, 2002) or on a combined measure (Borreggine & Kaschak, 2006). However, in studies using only sentences describing actions involving third-persons, the effect appears on the later YES-button press (Bergen & Wheeler, 2005; Tseng & Bergen, 2005). Since the stimuli in the current experiment all used only third-person arguments, it was anticipated that the effect would appear on the YES-button press, and not on the button release. All results reported below are therefore measures of YES-button press times.[4]

The Action-sentence Compatibility Effect involves faster button presses to indicate meaningfulness judgments when the direction in which participants have to move their hands is the same as the direction of motion implied by the sentence. We expected that if the progressive yields detailed mental simulation of event-internal actions, then this effect should be present in response to progressive sentences about concrete hand motions.

2.3. Results and discussion

No participants or items were deleted for reasons of accuracy or outlying mean response times. All trials with incorrect responses and all responses shorter than 50 ms or longer than 5000 ms were removed. This resulted in the exclusion of less than 4% of the data. There were three independent variables: Sentence-Direction (Towards or Away from the protagonist's body), Response-Direction (Towards or Away from the experimental participant's body), and Sentence-Type (Noun-manipulated or Verb-manipulated). An Action-sentence Compatibility Effect would appear on the interaction between Sentence-Direction and Response-Direction in the form of faster responses when the two directions matched than when they did not. This yielded the results reported in Table 1, presented graphically in Fig. 1.

We performed two three-way repeated-measures ANOVAs, one each with participants and items as random factors. These three-way analyses showed a significant interaction of Sentence-Direction with Response-Direction (the ACE), where manual responses were faster when they were in the same direction as the direction of motion described in the sentence. This was significant in both the participants analysis $F_1(1,69) = 9.82$, $p = 0.003$, $\eta_p^2 = 0.15$, and the items analysis $F_2(1,39) = 6.49$, $p = 0.02$, $\eta_p^2 = 0.15$. Only one other effect approached significance, which was the interaction of Sentence-Type (Noun-manipulated or Verb-manipulated) with Response-Direction; Toward responses were substantially slower than Away responses in Verb-manipulated sentences but not in Noun-manipulated sentences. This effect was marginally significant in the participants analysis $F_1(1,69) = 3.04$, $p = 0.09$, $\eta_p^2 = 0.06$, and significant in the items analysis $F_2(1,39) = 4.09$, $p = 0.05$, $\eta_p^2 = 0.10$. There was no interaction of the two different sentence types (4a versus 4b) with the Action-sentence Compatibility Effect.

Progressive sentences produced a reliable Action-sentence Compatibility Effect, regardless of whether the sentence encoded motion through contributions by the verb or object noun. This is in line with predictions made by simulation-based theories of language processing, which claim that the progressive signals the understander to mentally simulate the nucleus of a described event. The second experiment, below, tests whether the Action-sentence Compatibility Effect disappears when sentences are presented with perfect aspect.

3. Experiment 2: perfect

3.1. Method

The experiment was in almost all ways identical to Experiment 1. The major difference was in the aspect of stimuli, which all had perfect aspect. All sentences, including all criticals and fillers, were in the present perfect tense (3). Seventy members of the University of Hawaii community who had not participated in the first experiment took part in this experiment in exchange for course credit or $5. All incorrect responses and those that were faster than 50 ms or slower than 5000 ms were excluded from analysis. This resulted in elimination of less than 4% of the collected data.

We reasoned that if perfect sentences focus mental simulation on the endstate of an event, then we should find no significant effect of action-sentence compatibility on response times with these perfect sentences.

3.2. Results and discussion

The mean response times are shown in Table 2 and Fig. 2, below. A three-way repeated-measures ANOVA revealed only unpredicted and theoretically marginal effects. A main effect of Sentence-Type, where Verb-manipulated sentences led to slower responses than Noun-manipulated sentences, was significant by participants $F_1(1,69) = 4.70$, $p = 0.03$, $\eta_p^2 = 0.08$ but not by items $F_2(1,39) < 1$. An effect of Response-Direction was significant by both participants $F_1(1,69) = 4.8$, $p = 0.03$, $\eta_p^2 = 0.09$ and items $F_2(1,39) = 7.17$, $p = 0.01$, $\eta_p^2 = 0.12$, showing that Away responses were slower than Toward responses. There was also an interaction between Sentence-Type and Response-Direction, in which Away responses were much slower than Towards responses after Verb-manipulated sentences but not after Noun-manipulated sentences. This effect was significant by participants $F_1(1,54) = 5.05$, $p = 0.03$, $\eta_p^2 = 0.05$ but not items $F_2(1,39) < 1$.

Turning to the Action-sentence Compatibility Effect, responses were slightly slower when the action participants had to perform was incompatible with the action described by the sentence, but this interaction between Sentence-Direction and Response-Direction was not significant by either participants or items (both *F*s < 1). There were no other significant effects.

The absence of a significant Action-sentence Compatibility Effect when participants were presented with perfect sentences conforms to the notion that perfect aspect shuts off mental simulation of the core or nucleus of described events. Naturally, the absence of an effect in this case does not demonstrate that understanders perform absolutely no mental simulation of the core of described events when they are expressed using the perfect. But it does show that any mental simulation they do perform is different from that measured by Action-sentence Compatibility Effects reported else-

---

[4] Analyses of the button release measure in this study and in Experiment 2 showed no significant main or interaction effects.

Table 1
Results from Experiment 1: progressive sentences show an ACE.

Experiment 1: progressive sentences				
Sentence-Direction	Response-Direction	Mean RT (ms)	Std. deviation (ms)	*N*
Noun-manipulated				
Away	Away	355	112	70
Away	Toward	379	125	70
Toward	Away	380	162	70
Toward	Toward	358	109	70
Verb-manipulated				
Away	Away	358	112	70
Away	Toward	419	190	70
Toward	Away	360	134	70
Toward	Toward	362	104	70

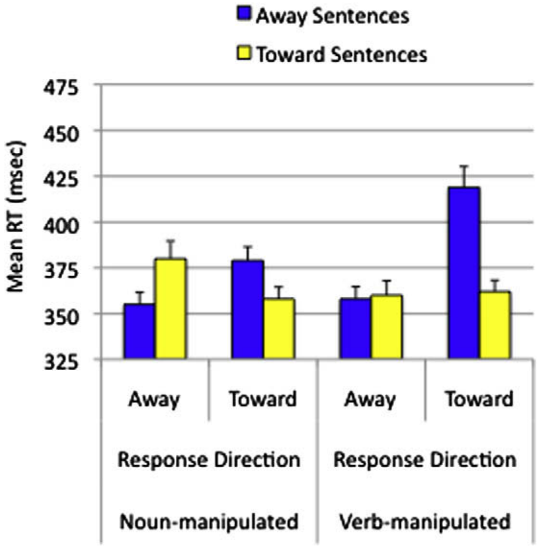

Fig. 1. Mean response times show a significant interaction between Sentence-Direction and Response-Direction for progressive sentences, regardless of sentence type. Error bars indicate standard error.

Table 2
Results from Experiment 2: Perfect sentences show no ACE.

Experiment 2: perfect sentences				
Sentence-Direction	Response-Direction	Mean RT (ms)	Std. deviation (ms)	*N*
Noun-manipulated				
Away	Away	410	234	70
Away	Toward	399	170	70
Toward	Away	391	168	70
Toward	Toward	384	142	70
Verb-manipulated				
Away	Away	425	188	70
Away	Toward	379	143	70
Toward	Away	452	265	70
Toward	Toward	394	158	70

where in the literature (e.g. Glenberg & Kaschak, 2002; Borreggine & Kaschak, 2006), and in Experiment 1, above.

4. Comparison across experiments

Experiment 1, which used progressive sentences, yielded a significant Action-sentence Compatibility Effect, showing that understanders processing progressive sentences construct mental

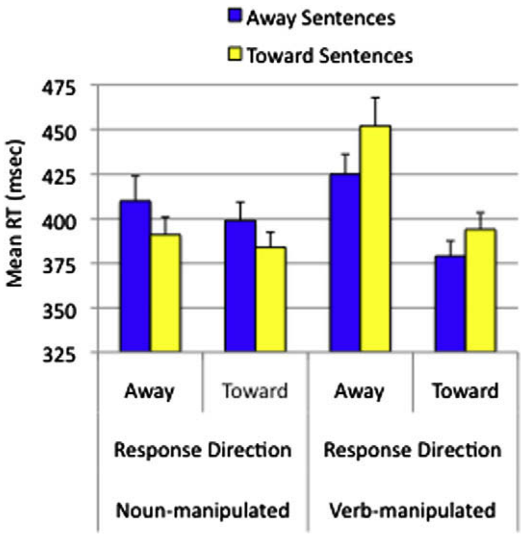

Fig. 2. Mean response times show no significant interaction between Sentence-Direction and Response-Direction for perfect sentences. Error bars indicate standard error.

simulations of the nuclei of described actions. However, Experiment 2, which used perfect sentences, showed no such effect, demonstrating that perfect sentences do not drive understanders to perform the same sort of mental simulation. To produce a more complete view of the data, we combined the results from these two experiments in a single analysis, and performed two 4-way repeated-measures ANOVAs, one each with participants and items as random factors. The four factors were Sentence-Direction, Response-Direction, Sentence-Type, and Aspect. These 4-way analyses showed only one main effect; Aspect was significant by items $F_2(1, 39) = 17.56$, $p < 0.0002$, $\eta_p^2 = 0.33$ but not by participants $F_2(1, 139) = 2.33$, $p = 0.13$, $\eta_p^2 = 0.04$; overall, sentences in the perfect experiment were processed more slowly than ones in the progressive experiment. We also found three unanticipated interaction effects. First, the interaction of Aspect with Response-Direction was significant in both the participants analysis $F_1(1, 139) = 5.96$, $p = 0.02$, $\eta_p^2 = 0.06$ and the items analysis $F_2(1, 39) = 6.61$, $p = 0.01$, $\eta_p^2 = 0.15$; Away responses were faster following Progressive sentences, while Toward responses were faster following Perfect sentences. Second, the three-way interaction among Sentence-Type, Sentence-Direction, and Aspect was significant by participants $F_1(1, 139) = 4.06$, $p = 0.05$, $\eta_p^2 = 0.04$ and nearly so by items $F_2(1, 39) = 3.81$, $p = 0.06$, $\eta_p^2 = 0.09$. And finally, there was a three-way interaction among Sentence-Type, Response-Direction, and Aspect, significant by both participants $F_1(1, 109) = 8.29$, $p < 0.005$, $\eta_p^2 = 0.07$ and items $F_2(1, 39) = 5.10$, $p = 0.03$, $\eta_p^2 = 0.12$. These last two effects each involved Sentence-Type, and as a result are hard to interpret. None of these unexpected effects involved the Action-sentence Compatibility Effect (Sentence-Direction by Response-Direction).

We now turn to the effects relevant to the matter under discussion in this paper. The Action-sentence Compatibility Effect, the interaction between Sentence-Direction and Response-Direction, yielded a significant effect in the analysis by participants $F_1(1, 139) = 3.98$, $p = 0.04$, $\eta_p^2 = 0.04$ and an effect approaching significance in the analysis by items $F_2(1, 39) = 2.80$, $p = 0.10$, $\eta_p^2 = 0.06$. Overall, responses were faster when their direction was compatible with that of the sentence. This weak compatibility effect understandably falls somewhere in between the effects in Experiments 1 (where there was a strong effect) and 2 (where there was none).

We were most interested in the effect of Aspect on the Action-sentence Compatibility Effect. The interaction among Sentence-Direction, Response-Direction, and Aspect approached significance in both the participants analysis $F_1(1, 139) = 3.11$, $p = 0.08$, $\eta_p^2 = 0.09$ and in the items analysis $F_2(1, 39) = 2.95$, $p = 0.09$, $\eta_p^2 = 0.08$. As we found from the separate results from the two experiments, Progressive sentences produced a large Action-sentence Compatibility Effect, while Perfect sentences did not. While these results do not meet the standard $p = 0.05$ threshold for significance, the relatively low p values for this three-way interaction ($p < 0.1$ by both participants and items) will be discussed below with reference to the hypothesis that while the perfect decreases mental simulation of the nucleus of a described event, it does not entirely block access to this mental simulation. No other main or interaction effects reached significance.

To summarize thus far, Experiment 1 (which used progressive stimuli) showed a strong Action-sentence Compatibility Effect, while this effect was absent in Experiment 2 (which focused on the perfect). In an ANOVA combining these experiments, the Action-sentence Compatibility Effect met significance. In addition, the interaction among Sentence-Direction, Response-Direction, and Aspect approached significance. The discussion below proposes several possible accounts of these findings.

5. General discussion

As expected, with progressive aspect in Experiment 1, we found an Action-sentence Compatibility Effect (faster manual responses when the Response-Direction was the same as the direction of motion described by the sentence). However, perfect sentences in Experiment 2 did not produce any such effect. These results (Fig. 3) can be straightforwardly explained in terms of effects of aspect on sentence processing. While progressive sentences drive understanders to mentally simulate the internal processes of described events, perfect sentences do not. This suggests that grammatical structures affect how language understanders engage their perceptual and motor systems to perform mental simulations of described content.

There has been relatively little previous work on how grammar affects language-driven mental simulation. The literature has convincingly shown through a variety of methods (Bergen, 2007) that sentences about perception and action yield modal mental imagery, and more specifically that this imagery is affected by content words like nouns and verbs. The first and more general of these points is made by the broad finding that sentences describing different events produce measurably different simulations (see, for instance, Zwaan et al., 2002; Glenberg & Kaschak, 2002; Richardson et al., 2003). For example, *The chair toppled* describes downwards motion of an object, and yields mental simulation in the lower part of the imagined visual field. By contrast, *The mule climbed* describes an upwards-oriented action, and yields perceptual simulation in the upper part of the visual field (Bergen et al., 2007). The second point is made by studies showing that manipulating specific content words (like nouns and verbs) across sentences influences the content of mental simulation (Zwaan et al., 2002; Richardson et al., 2003). For example, hand-motion verbs like "punch" demonstrably yield motor imagery revolving around moving the hand away from the body (Wheeler & Bergen, in press). These findings are in line with predictions made by simulation-based approaches to language understanding (Lakoff, 1987; Langacker, 1987; Talmy, 2000; Glenberg & Kaschak, 2002; Feldman & Narayanan, 2004; Bergen & Chang, 2005; MacWhinney, 2005; Feldman, 2006), which argue that content words provide the understander with specifications of what they should mentally simulate. On these views, words like nouns and verbs specify what category of object to simulate, what sort of event it is engaged in, what its properties are, and other specific instructions.

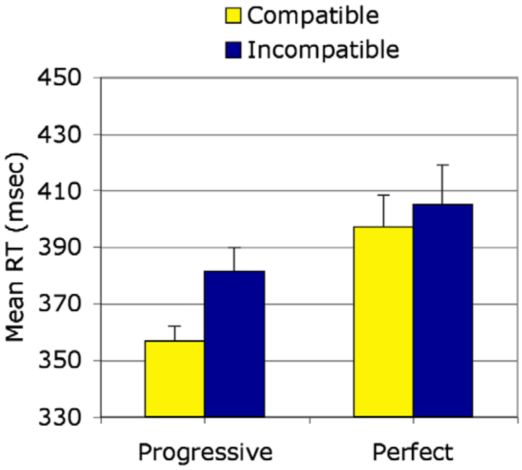

Fig. 3. Mean response time shows a reliable compatibility effect in the progressive but not the perfect. Whiskers represent standard error. The two Sentence-Types are conflated, and Sentence-Direction and Response-Direction are conflated into Compatibility.

But the work reported in this paper addresses a different claim of these same simulation-based theories. Several models of language understanding, most notably the Indexical Hypothesis (Glenberg & Kaschak, 2002) and Embodied Construction Grammar (Bergen & Chang, 2005) fundamentally depend on grammar combining constituents of an utterance and constraining their contributions to mental simulation. On such a view, grammar serves three functions with respect to mental simulation. First, like content words, it may contribute primary content to a mental simulation (Langacker, 1987; Goldberg, 1995; Kaschak & Glenberg, 2000). Second, it "assembles," "binds together," or "meshes" the contributions that content words like nouns and verbs make to the simulation (Glenberg & Kaschak, 2002; Bergen & Chang, 2005), so that the individual actors in a scene are simulated with the right roles. And third, the function tested in the work described in this paper, it modulates second-order properties of the mental simulation to be performed. We use the term "second-order" here because in this function, grammar serves not to directly propose content to be mentally simulated, but rather operates over this content. Given a scene to be simulated, grammar specifies what part to focus on (Chang et al., 1998), what perspective to adopt (MacWhinney, 2005; Brunyé, Ditman, Mahoney, Augustyn, & Taylor, 2009), what order to simulate the content in (Sato, Schafer, & Bergen, in preparation) and other such higher-order characteristics of simulation. The finding reported here, that progressive aspect drives understanders to mentally simulate the central process of a described motor event, while perfect aspect does not, corroborates the idea that grammar can yield such second-order effects.

The second-order effects of grammar on mental simulation are qualitatively different from the effects of content words. Whereas content words like *Mary, drawer*, or *open* point to specific experiences or categories of experience that can be simulated, grammatical aspect markers – and perhaps some other grammatical constructions – do not provide simulation content themselves. Rather, they appear to operate over the representations evoked by content words, modulating, for example, what part of an evoked simulation an understander focuses on, or the grain of detail with which the simulation is performed.

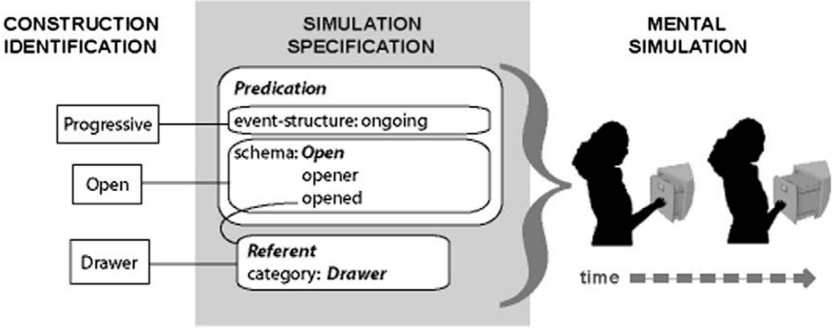

Fig. 4. Aspectual markers contribute to mental simulation through an intermediate representation;illustrated by a schematic representation of processing the verb phrase *is opening the drawer*.

While our results show that grammatical aspect affects the use of modality-specific systems, it is not clear from the results reported that its effects are entirely categorical. Although the results from the individual experiments were categorical (a significant Action-sentence Compatibility Effect for progressive sentences but none for perfect sentences), it could be that the underlying behavior is actually graded. Specifically, it could be that perfect sentences do not always fully shut off access to the nucleus of a described event. Some corroboration for this hypothesis comes from the merely marginal significance of the interaction of the ACE with Aspect in the combined analysis presented in Section 4. It could be that even perfect sentences allow subtle simulations of the nuclei of described actions – effects that our method was unable to detect at a conventional level of significance. Indeed, the theoretical literature leaves open the possibility that effects of aspect might not be categorical (Chang et al., 1998; Madden & Zwaan, 2003). We hope that future work will tease apart whether the effects of aspect – and other grammatical markers – are categorical or graded.

The modulating function of grammar has concrete consequences for theories of the language understanding faculty. The way an understander simulates an entity or event depends on second-order constraints provided by grammatical structures. There must therefore be room in any processing model for a means by which these second-order instructions produce their effects. But there are no simple solutions. For instance, it cannot be the case that as soon as a particular grammatical marker is encountered, the simulation process jumps into an appropriate state – for instance, simulating the nuclear details of a described action when a progressive marker is detected. The problems with this type of simple solution are twofold. First, grammatical markers do not necessarily temporally precede the words that indicate the content to be simulated. Markers of progressive aspect, for instance, straddle the verb that often specifies the action that is to be simulated. Second, the effects of each grammatical marker apply only to certain parts of an utterance, and a single utterance can have multiple such grammatical markers, each applying to potentially overlapping and non-contiguous subparts of the utterance.

We propose, therefore, that there must be a secondary process, running either in parallel with or following the parsing of words and grammar, which aligns the appropriate simulation contributions from the sentence with the second-order configurations that apply to them. On this account, the understanding process (grossly characterized for the phrase *is opening the drawer* in Fig. 4) would proceed from word and grammatical construction identification to mental simulation, but facilitated by a process assembling the contributions made by each (as argued by Feldman and Narayanan (2004) and Bergen and Chang (2005)). The outcome of this process, which has been called a *simulation specification* in the literature, include specifications of those components of perceptual and motor experience to be simulated, provided by the content words in the utterance. In the example in Fig. 4, the verb *open* specifies an 'Open' schema, which has two roles – for an 'opener' and an 'opened' – and the noun *drawer* identifies a referent that is of the 'Drawer' category. In addition, the simulation specification includes higher-order constraints on how the simulation is to be performed, provided by the grammatical markings. The Progressive construction in Fig. 4 specifies the opening event as having ongoing event structure, meaning that its internal process will be mentally simulated. (If the Perfect construction had been used instead of the Progressive, the internal process would have been represented as 'completed' instead of 'ongoing', and the resulting mental simulation would have included only the endstate of the action.) The simulation specification simultaneously serves as a means to combine the contributions to simulation from the various components of the utterance, and at the same time as a set of grounded pointers to components of mental simulation.

From a practical standpoint, uncovering the effects of particular grammatical markers on language-driven mental simulation informs our models of natural language acquisition and use. Human language would be unrecognizable without grammar, and aspect is one of the world's few universal linguistic traits. If the success of computational language understanding systems depends on the extent of their similarity to human language understanders, then incorporating mechanisms into them by which grammatical cues modulate simulation is critical to their success. From a broader perspective, understanding how grammar affects mental simulation is key to accounts of language and cognition. Grammar is often seen as the paragon of higher human cognitive functions because it is abstract, complex, and provides a number of uniquely human properties, like infinite combination, discussion of non-present events, and variable description of the same scene (Hockett, 1963). The findings reported here highlight the critical role that modal systems play in the organization of higher cognitive functions, by showing that grammar, a uniquely human and highly abstract cognitive capacity, hooks into pre-existing neural systems dedicated to perception and action.

Acknowledgments

Our thanks to Arthur Glenberg for invaluable consultation on this project, Kristin Ciano and Iris Kolberg for their help running subjects, members of the Cognitive Linguistics Research Group at the University of Hawai'i for feedback on the design, Ashley Maynard and Lori Yancura for useful comments on an early draft, and

Pia Knoeferle for helpful feedback on the submitted manuscript. All errors and omissions are our own.

Appendix A. Critical stimuli

Only progressive version is shown below. Perfect versions were identical except for aspect marking.

Away	Towards
Noun-differing pairs	
Rebecca is adjusting the thermostat	Lisa is adjusting her glasses
Pamela is beating the drum	Gregory is beating his chest
Dorothy is brushing her dog	Henry is brushing his hair
Shirley is brushing the couch	Virginia is brushing her teeth
Richard is cleaning the wall	Jean is cleaning her ear
Ben is feeding his child	Robert is feeding himself
Donna is grabbing her keys	Louis is grabbing his nose
Melissa is grabbing the doorknob	Roy is grabbing his ear
Amy is lighting the grill	Willie is lighting his cigarette
Chris is patting the cat	Larry is patting his tummy
Albert is pinching the baby	Brian is pinching his chin
Terry is pushing the elevator button	Thomas is pushing his belly button
Juan is putting in his favorite CD	Fred is putting in his contact lens
Mary is rubbing the magic lamp	Jerry is rubbing his eye
Catherine is scratching the cat	Kelly is scratching her head
Mildred is squeezing the mustard bottle	Barbara is squeezing the back of her neck
Elizabeth is tucking in the sheets	Jonathan is tucking in his shirt
Eric is washing his desk	Joan is washing her face
William is washing the window	Steve is washing his hair
Helen is wiping the counter	Brenda is wiping her mouth
Verb-differing pairs	
Beverly is closing the drawer	Nicholas is opening the drawer
Judith is closing the cupboard	Teresa is opening the cupboard
Patricia is displaying her ring	Janice is snatching the ring
Julie is flicking a cigarette	Harry is smoking a cigarette
Carl is flipping the burger	Karen is eating a burger
Kimberly is hanging up the phone	Arthur is answering the phone
Betty is pushing the door	Cheryl is pulling the door
Angela is putting down the toys	Dennis is picking up the toys
Lawrence is roasting a marshmallow	Joyce is stealing a marshmallow
Judy is rubbing the dog's belly	Stephanie is rubbing her belly
Theresa is showing off her fingernails	Donald is biting his fingernails
Maria is spitting out the water	Margaret is swallowing the water
Ashley is stretching her arms	Michelle is crossing her arms
Carol is taking off her glasses	Evelyn is putting on her glasses
George is taking off the jacket	Brandon is putting on the jacket
Carolyn is throwing away the earplugs	Edward is putting in the earplugs
Kathleen is throwing the pie	James is eating the pie
Bruce is tossing out the water	Matthew is taking a sip of water
Harold is tossing back a fish	Stephen is dragging in a fish
Joshua is tossing a Q-tip	Ryan is using a Q-tip

References

Aziz-Zadeh, L., Wilson, S., Rizzolatti, G., & Iacoboni, M. (2006). A comparison of premotor areas activated by action observation and action phrases. *Current Biology, 16*(18), 1818–1823.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–609.

Bencini, G., & Goldberg, A. (2000). The contribution of argument structure constructions to sentence meaning. *Journal of Memory and Language, 43*, 640–651.

Bergen, B. (2007). Experimental methods for simulation semantics. In Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson, & Michael J. Spivey (Eds.), *Methods in cognitive linguistics* (pp. 277–301). John Benjamins.

Bergen, B., & Chang, N. (2005). Embodied construction grammar in simulation-based language understanding. In J.-O. Östman & M. Fried (Eds.), *Construction grammar(s): Cognitive grounding and theoretical extensions.* (pp. 147–190). Amsterdam: Benjamin.

Bergen, B., & Wheeler, K. (2005). Sentence understanding engages motor processes. In *Proceedings of the twenty-seventh annual meeting of the cognitive science society*.

Bergen, B., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*.

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

Borreggine, K. L., & Kaschak, M. P. ((2006). The action-sentence compatibility effect: Its all in the timing. *Cognitive Science, 30*, 1097–1112.

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*, 27–32.

Carreiras, M., Carriedo, N., Alonso, M. A., & Fernández, A. (1997). The role of verb tense and verb aspect in the foregrounding of information during reading. *Memory and Cognition, 25*, 438–446.

Chang, N. (2004). *A computational model of comprehension-based construction acquisition*. Stanford, CA: Child Language Research Forum.

Chang, N., Gildea, D., & Narayanan, S. (1998). A dynamic model of aspectual composition. In *Proceedings of the twentieth annual conference of the cognitive science society*.

Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review, 112*(2).

Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.

Denis, M., & Kosslyn, S. M. (1999). Scanning visual mental images: A window on the mind. *Current Psychol Cognition, 18*, 409–465.

Dowty, D. (1977). Toward a semantic analysis of verb aspect and the English "imperfective" progressive. *Linguistics and Philosophy, 1*, 45–77.

Ehrsson, H. H., Geyer, S., & Naito, E. (2003). Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part specific motor representations. *Journal of Neurophysiology, 90*, 3304–3316.

Feldman, J., & Narayanan, S. (2004). Embodied meaning in a neural theory of language. *Brain and Language, 89*, 385–392.

Feldman, J. (2006). *From molecule to metaphor: A neural theory of language.* Cambridge: MIT Press.

Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 182–196.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory–motor system in reason and language. *Cognitive Neuropsychology, 22*, 455–479.

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

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure.* Chicago: University of Chicago Press.

Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex, *Neuron, 41*(2).

Hockett, C. (1963). The problem of universals in language. In J. Greenberg (Ed.), *Universals of language*. Cambridge, MA: MIT Press.

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

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M. E., & Blanchard, A. A. (2005). Perception of motion affects language processing. *Cognition, 94*, B79–B89.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience, 2*, 635–642.

Lakoff, G. (1987). *Women, fire and dangerous things*. Chicago: University of Chicago Press.

Langacker, R. W. (1987). *Foundations of cognitive grammar.* Stanford University Press.

Lotze, M., Montoya, P., Erb, M., Hülsmann, E., Flor, H., Klose, U., et al. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: An fMRI study. *Journal of Cognitive Neuroscience, 11*(5), 491–501.

MacWhinney, B. (2005). The emergence of grammar from perspective taking. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory language, and thinking* (pp. 198–223). Cambridge, England: Cambridge University Press.

Madden, C., & Therriault, D. (2009). Verb aspect and perceptual simulations. *The Quarterly Journal of Experimental Psychology, 62*(7), 1294–1302.

Madden, C. J., & Zwaan, R. A. (2003). How does verb aspect constrain event representations? *Memory and Cognition, 31*, 663–672.

Magliano, J. P., & Schleich, M. C. (2000). Verb aspect and situation models. *Discourse Processes, 29*, 83–112.

Mellet, E., Bricogne, S., Crivello, F., Mazoyer, B., Denis, M., & Tzourio-Mazoyer, N. (2002). Neural basis of mental scanning of a topographic representation built from a text. *Cerebral Cortex, 12*(12), 1322–1330.

Narayanan, S. (1997). *KARMA: Knowledge-based active representations for metaphor and aspect*. PhD dissertation, Computer Science Division, University of California, Berkeley, USA.

Porro, C. A., Francescato, M. P., Cettolo, V., Diamond, M. E., Baraldi, P., Zuian, C., et al. (1996). Primary motor and sensory cortex activation during motor performance and motor imagery: A functional magnetic resonance imaging study. *Journal of Neuroscience, 16*, 7688–7698.

Pulvermüller, F., Haerle, M., & Hummel, F. (2001). Walking or talking?: Behavioral and neurophysiological correlates of action verb processing. *Brain and Language, 78*, 143–168.

Pulvermuller, F. (2003). *The neuroscience of language. On brain circuits of words and serial order*. Cambridge University Press.

Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science, 27*, 767–780.

Sato, M., Mengarelli, M., Riggio, L., Gallese, V., & Buccino, G. (2008). Task related modulation of the motor system during language processing. *Brain and Language, 105*(2), 83–90.

Sato, M., Schafer, A., & Bergen, B. (in preparation). *Incremental construction of mental images in Japanese sentence comprehension*.

Talmy, L. (2000). *Toward a cognitive semantics*. Cambridge: MIT.

Tettamanti, M., Buccino, G., Saccuman, M. C., 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.

Tseng, M., & Bergen, B. (2005). Lexical processing drives motor simulation. In *Proceedings of the twenty-seventh annual conference of the cognitive science society*.

Wheeler, K., & Bergen, B. (in press). Meaning in the palm of your hand. In Sally Rice & John Newman (Eds.), *Empirical and experimental methods in conceptual structure, discourse, and language*. Stanford: CSLI.

Willems, R. M., & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture and action: A review. *Brain and Language, 101*(3), 278–298.

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