

# The activation of modality-specific representations during discourse processing



Christopher A. Kurby<sup>a,\*</sup>, Jeffrey M. Zacks<sup>b</sup>

<sup>a</sup> Grand Valley State University, Department of Psychology, 2224 Au Sable Hall, Allendale, MI 49401, United States

<sup>b</sup> Washington University in St. Louis, Department of Psychology, One Brookings Drive, St. Louis, MO 63130, United States

## ARTICLE INFO

### Article history:

Accepted 7 July 2013

Available online 8 August 2013

### Keywords:

Perceptual simulation

Imagery

Language comprehension

Discourse comprehension

Neuroimaging

## ABSTRACT

Previous research has shown that readers generate mental images of events. Most studies have investigated imagery during the reading of short texts, which also included explicit judgment tasks. In two fMRI studies, we assessed whether modality-specific imagery occurs during naturalistic, *discourse* comprehension. We identified clauses in the texts that elicited auditory, motor, or visual imagery. In both studies, reading motor imagery clauses was associated with increases in activity in left postcentral and precentral sulci, and reading auditory imagery clauses was associated with increases in left superior temporal gyrus and perisylvian language-related regions. Study 2 compared presentation of connected discourse to a condition in which unconnected sentences were presented, preventing the establishment of global coherence. Sensorimotor imagery was strongest when readers were able to generate a globally coherent discourse representation. Overall, these results suggest that modality-specific imagery occurs during discourse comprehension and it is dependent on the development of discourse-level representations.

© 2013 Elsevier Inc. All rights reserved.

## 1. Introduction

Events in narrative texts are typically described in ways that imply how things sound, look, and feel. Corresponding with this, people often report rich subjective experiences reflecting these features when reading stories (Mar & Oatley, 2008). There is good reason to think that representing the perceptual and motor features of described events is an important component of comprehension (Barsalou, 1999; Glenberg, 1997; Zwaan, 2004; Zwaan & Radvansky, 1998). What is the format of perceptual and motor information that is used for comprehension? This is currently a matter of debate.

One way of describing this question is in terms of *situation models*. A situation model is a representation of the events described by a discourse (Johnson-Laird, 1983; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). Readers construct situation models, in part, by activating relevant knowledge associated with the event, and combining it in ways consistent with how the activity is described (Zwaan & Radvansky, 1998). Sensorimotor simulation theories argue that the representational format of situation models is in large part sensory and motor: Situation models are implemented by the same sensorimotor neural representations formed while physically perceiving the event (Barsalou, 1999, 2008; Glenberg, 1997; Hesslow, 2002). Thus, comprehenders simulate perceptual

and motor properties of experience as a normal part of ongoing comprehension (Zwaan, 2004). These simulations are mental reenactments of the sensorimotor states described by the events in the text. Though most simulations are thought to be unconscious reenactments, the related phenomenon of mental imagery is proposed to be a conscious form of simulation (Barsalou, 2008).

An alternative view of situation model construction is that comprehenders rely on symbolic and abstract representations that are functionally separate from the perceptual systems (Anderson & Lebiere, 1998; Fodor, 1975; Mahon & Caramazza, 2008; Markman & Brendl, 2005; Pylyshyn, 1981, 2002). Kintsch (1998) proposed that comprehenders generate amodal propositional representations that connect text content to world knowledge and inferences. According to this view, comprehension proceeds largely without the need to ground representations to perceptual referents or images. Recent views of comprehension propose a more complex relationship between perceptual and amodal representations during the construction of situation models. Louwerse (2008) proposed that readers rely differentially on amodal or sensorimotor representations depending on task and goals (for evidence of this view, see Louwerse & Jeuniaux, 2010). Mahon and Caramazza (2008) argued that the conceptual system might be both embodied and abstract and that simulations may be downstream from the activation of higher-level abstract representations. On their grounding by interaction hypothesis, conceptual knowledge is inherently abstract but is not complete without the perceptual and motor knowledge that links it to the physical world. According

\* Corresponding author.

E-mail address: [kurbyc@gvsu.edu](mailto:kurbyc@gvsu.edu) (C.A. Kurby).

to their view, the activation of a concept centrally retrieves its abstract content and simulations may follow in order to elaborate on, or complement, the abstract representations constructed during reading.

There is a large amount of support for the sensorimotor simulation view of comprehension (see the following for reviews: Barsalou, 2008; Fischer & Zwaan, 2008; Gallese & Lakoff, 2005; Glenberg & Gallese, 2012; Kemmerer & Gonzalez-Castillo, 2010). Those studies have shown that readers activate modality-specific perceptual representations during the comprehension of text. When reading isolated verbs (Hauk, Johnsrude, & Pulvermüller, 2004; Willems, Hagoort, & Casasanto, 2010) and isolated action phrases (Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006; Tettamanti et al., 2005) somatosensory and motor cortex shows topographically organized activation corresponding to the effector associated with the verb. (However, see Postle, McMahon, Ashton, Meredith, & de Zubicaray, 2008.) The comprehension of verbs and action phrases has also been associated with the activation of premotor cortex (Tettamanti et al., 2005), left postcentral sulcus, left anterior supramarginal gyrus, left fusiform gyrus, and posterior inferior temporal gyrus (Desai, Binder, Conant, & Seidenberg, 2010). Transcranial magnetic stimulation (TMS) studies have demonstrated that somatomotor regions may be critical in the representation of action knowledge about verbs by selectively speeding their recognition by limb-location (Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005; Willems, Labruna, D'Esposito, Ivry, & Casasanto, 2011). Behavioral studies show that reading about actions affects the execution of actions (Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006).

There is also evidence for the activation of sensorimotor representations in the visual domain. Research has shown that participants activate a portion of left fusiform gyrus, a region shown to respond during visual object recognition (Kanwisher, Downing, Epstein, & Kourtzi, 2001; Orban, Van Essen, & Vanduffel, 2004) and during explicit visual imagery tasks (D'Esposito et al., 1997; Thompson-Schill, Aguirre, D'Esposito, & Farah, 1999), when verifying visual properties of objects (cat – whiskers), and the color of objects (banana – yellow; Simmons et al., 2007). Wheeler, Petersen, and Buckner (2000) also found activation of left fusiform gyrus during the recall of pictures from verbal labels. Desai et al. (2010) found activation of posterior superior temporal sulcus (pSTS), a region thought to process visual biological motion (Allison, Puce, & McCarthy, 2000; Grossman et al., 2000), when comprehending phrases implying visual content (e.g., “I inspect the toy.”). Behavioral results provide further evidence for the activation of visual representations during language tasks (Fischer & Zwaan, 2008). In these studies, participants made judgments about pictures in ways that implied that they were simulating aspects of those pictures during language comprehension (Stanfield & Zwaan, 2001; Zwaan, Madden, Yaxley, & Aveyard, 2004; Zwaan, Stanfield, & Yaxley, 2002).

Some studies have shown that people activate auditory representations for tasks that access sound representations (for a review see Hubbard, 2010). For example, studies have found that making judgments about the sounds of objects, or recalling their sound, is associated with the activation of auditory regions such as left posterior superior temporal gyrus (Kellenbach, Brett, & Patterson, 2001; Wheeler et al., 2000), middle temporal gyrus (Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008), and bilateral inferior frontal gyrus (Wheeler et al., 2000). Yao, Belin, and Scheepers (2011) found that the silent reading of direct speech activated voice-selective brain regions in temporal and frontal cortex. These studies show that the processing of sounds, both environmental and speech, in text and imagery engage secondary auditory cortex, and may also engage regions associated with perisylvian language areas, such as the inferior frontal gyrus. Recent behavioral evidence

shows that readers engage in auditory imagery when reading sentences that imply sounds (Brunyé, Ditman, Mahoney, Walters, & Taylor, 2010) or depict dialog (Kurby, Magliano, & Rapp, 2009).

Although this body of literature supports the possibility that readers activate modality-specific sensorimotor representations during reading, evidence also suggests that language processing does not always recruit such representations. For example, consider the *spatial iconicity effect*—the finding that semantic judgments for word pairs presented in a spatial arrangement congruent with their real-world arrangement (e.g., “roof” above “basement”) are faster than when presented in an incongruent arrangement (e.g., “basement” above “roof”; Zwaan & Yaxley, 2003). Louwerse (2008) found that this effect can be explained by linguistic properties of the text (e.g., word order). Louwerse and Jeuniaux (2010) found that the type of task (visual vs. verbal) and materials (visual vs. verbal) influence whether readers rely on perceptual or abstract representations. Some recent neuroimaging research has shown that readers also rely on modality-independent representations during verb and action word comprehension. Bedny, Caramazza, Grossman, Pascual-Leone, and Saxe (2008) investigated the response selectivity of a region shown to increase in activity during action verb comprehension: posterior lateral temporal cortex, which has been postulated to represent visual-motion information about actions (Damasio et al., 2001; Kable & Chatterjee, 2006). Bedny et al. (2008) found that this region increased in activity when participants made judgments about verbs, but that this activity was not modulated by amount of motion implied by the verb, and did not overlap with regions identified by biological and non-biological motion localizers. Bedny et al. (2008) concluded that activity in this region might instead reflect the activation of non-perceptual categorical information about events.

In sum, the available behavioral and neurophysiological evidence supporting sensorimotor simulation views shows that the activation of modality-specific representations can be detected, when it occurs, and suggests that humans can use such representations to perform language-related tasks (but perhaps not necessarily so). However, it leaves open the critical question of whether the activation of these representations occurs during normal discourse comprehension. The behavioral and the neurophysiological studies to date share two important limitations: First, they have all used single words, single sentences, or very short artificial texts as stimulus materials. Generalizing from such “textoids” to larger discourse comprehension is difficult because these short texts are unlikely to capture the global coherence building and maintenance processes necessary for discourse comprehension (Graesser, Millis, & Zwaan, 1997). Second, most studies have used concurrent judgment tasks – word/property/sentence verification, sensibility judgment tasks, perceptual judgment tasks, motor movement tasks, etc. – that likely alter normal comprehension processes and may promote the activation of sensorimotor representations (Graesser et al., 1997; Louwerse, 2008). This means that it is unclear to what extent task demands influence sensorimotor activation during language comprehension, and discourse comprehension specifically (Louwerse, 2008). The main goal of the current study was to investigate whether the activation of modality-specific representation occurs during discourse comprehension in reading conditions that do not contain demands beyond what is typical in normal comprehension: When reading extended narratives with no task other than to comprehend.

In this study, we reanalyzed data from two previous studies (Speer, Zacks, & Reynolds, 2007; Yarkoni, Speer, & Zacks, 2008) in which participants read extended narratives while their brain activity was recorded with fMRI. The participants' only task was to read for comprehension. They did not engage in any judgment tasks during reading and were not given special reading goals.

Through norming and coding procedures, we identified clauses in the texts that elicited auditory imagery (e.g., descriptions of sounds), visual imagery (e.g., descriptions of visual scenes), and motor imagery (e.g., descriptions of actions), and used this coding to predict brain activity during reading. If readers activate modality-specific representations during discourse comprehension, then they should activate sensorimotor brain regions that correspond with specific imagery modalities when processing implied perceptual information in the text. Specifically, reading text units that imply motor actions should be associated with the activation of premotor cortex, postcentral sulcus, and possibly regions in primary motor and somatosensory cortex. Reading about auditory information should be associated with the activation of superior temporal gyrus, middle temporal gyrus, and regions in perisylvian language-related areas. Finally, reading about visual information should be associated with activation in fusiform gyrus and other secondary visual regions, such as pSTS.

## 2. Study 1

In Study 1, we asked whether readers activate modality-specific brain regions during discourse comprehension, without explicit judgment tasks. To do so, we re-analyzed data from Speer et al. (2007). The neuroimaging methods have previously been described by Speer et al. (2007) and Speer, Reynolds, Swallow, and Zacks (2009). We will briefly summarize those methods here, and will elaborate on new norming and analysis that was performed for the current study.

### 2.1. Method

#### 2.1.1. Neuroimaging methods

Twenty-eight right-handed individuals (ages 19–34 years, 20 women) read four narratives while their brain activity was recorded with fMRI. The narratives were scenes from the book *One Boy's Day* (Barker & Wright, 1951), which is an observational record of the everyday activities of a seven-year-old boy named Raymond. The four scenes described Raymond getting up and eating breakfast ("Waking up"; 1368 words, 192 clauses), having an English lesson ("Class work"; 1182 words, 172 clauses), having a music lesson ("Music lesson"; 1404 words, 215 clauses), and playing with his friends in the schoolyard ("Play before school"; 1104 words, 178 clauses). The stimuli can be downloaded from <http://dcl.wustl.edu/DCL/stimuli.html> (follow the links for the Speer et al., (2007) stimuli).

Participants read each narrative one word at a time. (A number of studies have shown that readers engage in situation model construction when reading word-by-word (e.g., Friesse, Rutschmann, Raabe, & Schmalhofer, 2008).) Each word was displayed for 200 ms followed by a 150 ms/syllable blank delay. This was done in order to control for eye-movements and to have a precise timing of event onsets. Each text was presented in a separate run, ranging in time from 8.5 to 10.9 min. The order of narratives was counter-balanced across participants. Readers completed four-alternative forced-choice comprehension tests after each passage to verify that they understood the texts ( $M = 82.74\%$  correct; see Friesse et al., 2008).

Images were acquired on a 3-T Siemens Vision MRI scanner (Erlangen, Germany). For data registration, functional images ( $T2^*$ -weighted images with 32 slices,  $4.0 \times 4.0$  mm in-plane resolution, acquired every 2.048 s) were aligned with structural images from a  $T2$ -weighted fast turbo spin-echo scan, which were then aligned to high-resolution  $T1$ -weighted images. The functional data were preprocessed following standard protocols, including correction of timing offsets with spline interpolation, data

alignment, and normalization of image intensity (see Speer et al., 2007 for details), and warped to a standard stereotactic space with 3 mm isotropic voxels (Talairach & Tournoux, 1988).

#### 2.1.2. Collection of imagery norms

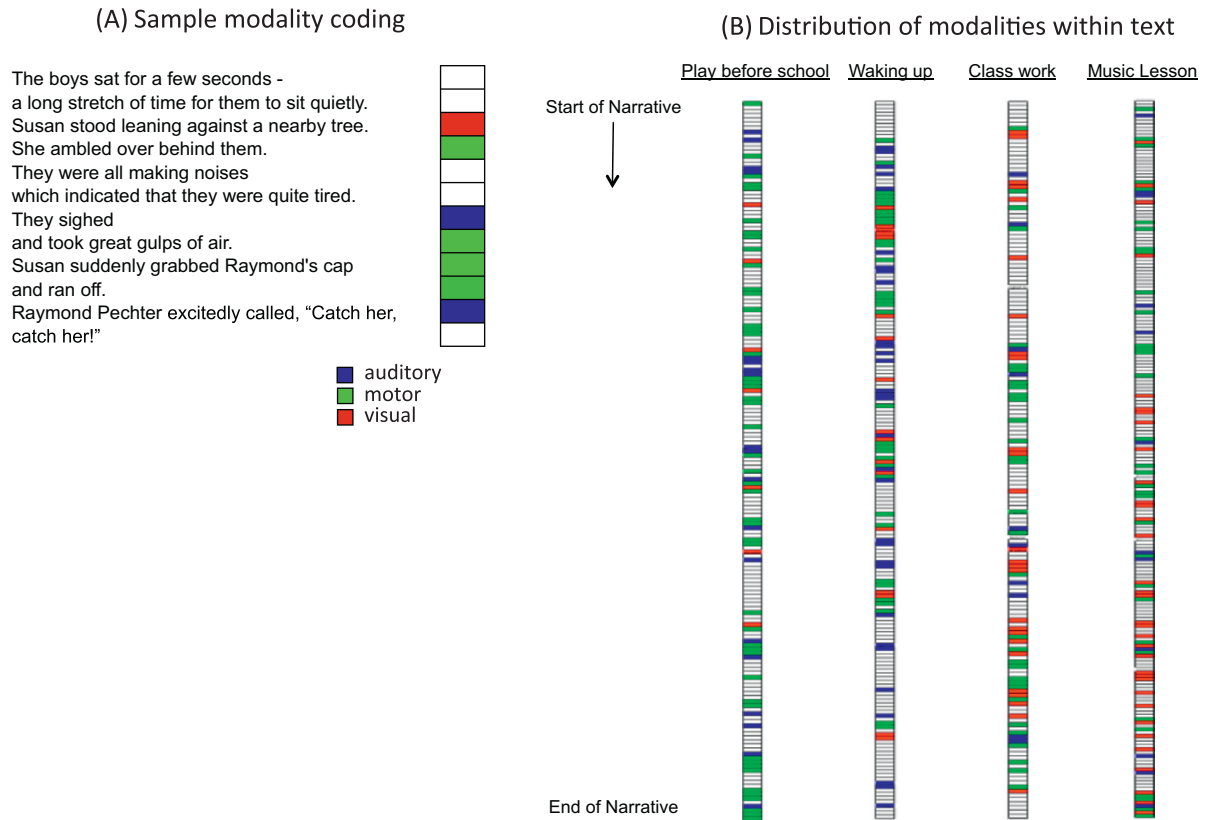
We conducted a norming study to identify high-imagery clauses in the narratives. We reasoned that participants would be most likely to activate sensorimotor representations at these clauses during silent reading in the scanning session. A set of 33 participants, who did not participate in the scanning session, read each text one clause at a time on a computer screen. After reading each clause, participants rated the vividness of the imagery they experienced while reading that clause on a scale from 1 (no imagery at all) to 6 (as vivid as perception). The imagery scale was adapted from the scale used in the Vividness in Visual Imagery Questionnaire (VVIQ; Marks, 1995). Participants were informed that we were interested in any type of imagery they experienced, including how something looks, sounds, smells, feels, or tastes, and also imagery for what it feels like to perform an action depicted in the story. Participants were not asked to classify what kind of imagery they experienced—simply to rate its vividness. Once an imagery rating was provided, the clause was removed and the next clause was presented. Participants rated all clauses of all four narratives; the order of narratives was counterbalanced across participants. The mean clause imagery rating was 3.47 ( $SD = 0.46$ ).

#### 2.1.3. Imagery modality coding

To code for imagery modality, we first computed the mean imagery rating for each clause. Second, for each text, clauses in the upper two quintiles of imagery strength were selected for imagery modality coding. We coded for whether these high-imagery clauses elicited auditory, motor, or visual imagery, blind to the fMRI data. We coded for these imagery modalities rather than asking participants for ratings because we wanted to ensure that there were clear and definable differences between modality types and it was unclear whether participants have conscious access to the types of imagery they may be experiencing at any moment. Fig. 1a presents an example of the modality coding, and Fig. 1b shows the distribution of imagery modality type within each text. *Auditory imagery* clauses implied or explicitly described a sound (e.g., "and giggled in a breathless way"), or were a line of dialog. *Visual imagery* clauses described the visual appearance of a scene or a visual characteristic of an object/character (e.g., "He looked quite cocky with his legs spread apart and his arms akimbo."). *Motor imagery* clauses described or implied a character executing an action (e.g., "Raymond picked up his hat"). In total, there were 72 auditory imagery clauses (38 clauses contained dialog, and 34 described non-dialog sounds),<sup>1</sup> 84 visual imagery clauses, and 149 motor imagery clauses.<sup>2</sup> (We did not code for secondary modalities that could be imaged from the clause, or potential multi-modal imagery. In preliminary discussions, we discovered that it was difficult to obtain acceptable inter-rater agreement on such coding. Additionally, the number of observations per condition on such categories would be too low to provide adequate statistical power.) The remaining 451 clauses were classified as low-imagery and served as the comparison condition for analysis. Cohen's kappa across the

<sup>1</sup> Exploratory analyses investigating differences between these two types of auditory clauses indicated little difference in brain response to them. Given that there is a small amount of items in these sub-categories, and that the study was not specifically designed to test any such differences, we collapsed across these two item types.

<sup>2</sup> We also coded for haptic, olfactory, and gustatory imagery. However, there were not enough clauses implying these imagery types (haptic = 0 clauses; olfactory = 1 clause; gustatory = 1 clause) to include in the analyses.



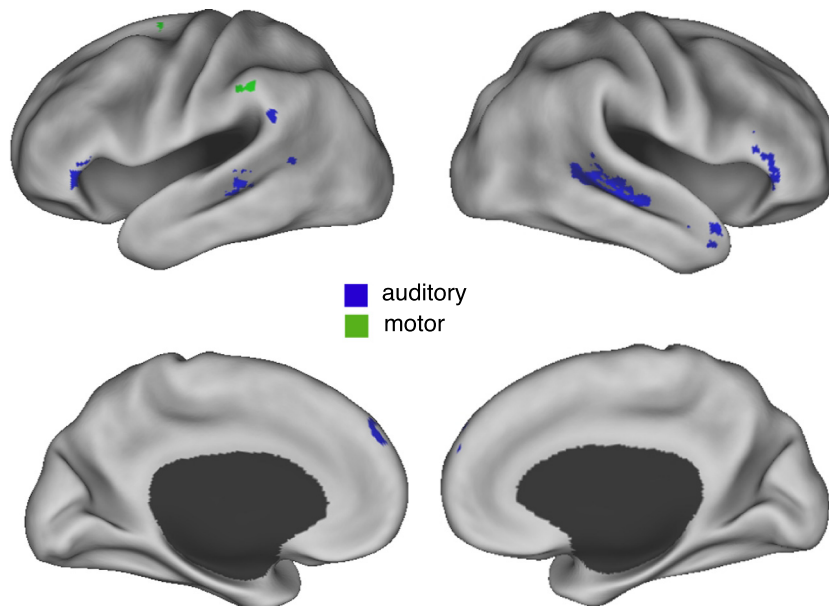
**Fig. 1.** (A) Example modality coding. The colored column on the right indicates the imagery modality assignment. The white cells indicate that the clause is a low-imagery clause. (B) Distribution of imagery modalities within each text.

three imagery modalities was .76. Disagreements were resolved by a discussion blind to the fMRI data.

#### 2.1.4. Analysis of the neuroimaging data

For each participant, general linear models (GLMs) were computed to estimate the brain response to each imagery modality type. We treated individual clauses as trials in a rapid-event-

related data analysis. The interval between successive instances of each clause type varied considerably, which made it possible to accurately estimate independent effects for each imagery type (Zacks et al., 2001). Four predictor variables were created to code the clause onset of each of the three imagery modality types and the low-imagery control. To generate regressors for the GLMs, all four clause types were each coded as a brief pulse of activity at



**Fig. 2.** Regions showing modality-specific imagery effects in Study 1.



**Table 1**

Regions from Study 1 that increased in activity for the different imagery modalities, and their effects in Study 2.

Imagery modality	Center of mass coordinates			Hemisphere	Region	Volume (mm <sup>3</sup> )	Peak z statistic	Study 2 scrambled	Study 2 story	Study 2 imagery X coherence interaction
	x	y	z							
Auditory	–61	–54	27	Left	Posterior superior temporal gyrus, at temporo-parietal junction (BA 22/40)	108	4.70	–	✓	–
	–60	–46	23	Left	Superior temporal gyrus (BA 22)	189	4.96	–	✓	–
	–55	–55	10	Left	Superior temporal sulcus (BA 22)	405	5.20	–	✓	✓
	–53	–35	–3	Left	Middle temporal gyrus (BA 22)	513	5.19	–	✓	✓
	–48	26	–1	Left	Inferior frontal gyrus (BA 45/47)	864	5.65	–	✓	✓
	–5	52	37	Left	Medial superior frontal gyrus (BA 9)	621	5.59	–	✓	✓
	8	56	27	Right	Medial superior frontal gyrus (BA 9)	54	4.99	–	–	–
	11	55	37	Right	Superior frontal gyrus (BA 9)	81	5.70	–	✓	–
	49	11	–22	Right	Anterior superior temporal gyrus (BA 38)	324	5.69	–	✓	–
	52	–35	–1	Right	Superior temporal sulcus (BA 22)	3105	5.75	–	✓	✓
	53	28	5	Right	Inferior frontal gyrus (BA 47)	648	5.43	–	✓	✓
	53	3	–18	Right	Superior temporal sulcus (BA 22)	81	5.02	–	✓	–
	54	24	23	Right	Middle frontal gyrus (BA 46)	162	5.15	–	–	–
Motor	–62	–35	38	Left	Postcentral sulcus (BA 5/40)	1026	6.23	–	✓	✓
	–24	–2	59	Left	Precentral sulcus (BA 6)	135	4.62	–	✓	✓
Visual	–	–	–	–	–	–	–	–	–	–

Note: Coordinates listed are the center of mass coordinates in Talairach space. Check marks indicate whether the region showed a significant effect in Study 2 for the scrambled condition, story condition, and a imagery X condition interaction, corrected for multiple comparisons using the Bonferroni method (13 regions for auditory imagery and 2 regions for motor imagery).

the beginning of the clause, which was convolved with a canonical hemodynamic response function (Boynton, Engel, Glover, & Heeger, 1996) to generate a predicted fMRI response. Ten additional regressors were included to code for effects of no interest: terminal and non-terminal punctuation, differences between runs, and linear trend within each run.

Paired-sample *t* tests were conducted comparing the effect estimates of each type of high-imagery variable (auditory, motor, visual) with the effect estimates for the low-imagery variable (Auditory – LowImagery; Motor – LowImagery; Visual – LowImagery), using subject as a random effect. Maps of *t*-statistics were generated and converted to *z*-statistic maps, which were thresholded to control for a map-wise false positive rate at .05 by retaining clusters of at least 2 voxels with *z* values greater than 4.5 (McAvoy, Ollinger, & Buckner, 2001). These were projected onto an inflated cortical surface for visualization, using CARET with the PALS atlas (Van Essen, 2005; Van Essen et al., 2001).

## 2.2. Results and discussion

The regions that showed increases in activation for the different imagery modalities, compared to low-imagery clauses,<sup>3</sup> are illustrated in Fig. 2 and listed in Table 1. Imagery-related activation occurred in or near modality-specific brain regions.

The reading of auditory imagery clauses was associated with activation in a number of regions in or near auditory cortex (BA 22), notably left superior temporal gyrus and bilateral superior temporal sulcus. Auditory imagery was also associated with activation in perisylvian language-related regions: at the junction of the left superior temporal gyrus and supramarginal gyrus adjacent to the posterior end of the sylvian fissure (BA, 22/40) and the left inferior frontal gyrus (BA 44/45). In the left hemisphere, these two regions correspond approximately to the classical Wernicke's and Broca's areas (e.g., Catani, Jones, & ffytche, 2005; Ferstl, Neumann,

Bogler, & von Cramon, 2008). In most areas activity was bilateral and approximately equally strong in the left and right hemispheres. (One exception was the most superior/posterior extent of the activity in the left superior temporal gyrus, at the temporo-parietal junction. This was not present in the right hemisphere.)

Reading motor imagery clauses was associated with activation in left premotor (BA 6) and left secondary sensorimotor cortex (BA 5/40), as can be seen in Fig. 2 and Table 1. These regions are selectively associated with performing grasping motions and with touch sensation on the hand (Castiello, 2005; Porro, Francescato, Cettolo, Diamond, et al., 1996) and are in close proximity to regions found to respond to the reading of action verbs (Aziz-Zadeh et al., 2006; Hauk et al., 2004; Kemmerer & Gonzalez-Castillo, 2010; Tettamanti et al., 2005; Willems, Hagoort, et al., 2010; Willems, Toni, Hagoort, & Casasanto, 2010).

There were no significant imagery effects associated with reading visual imagery clauses.

In short, reading clauses high in auditory or motor imagery led to increases in activity in modality-specific brain regions. The activation of regions in and near auditory cortex and language-related processing areas supports the possibility that readers activate auditory representations when reading about sounds and dialog. Related to this, the activation in superior frontal gyrus and TPJ are consistent with work showing that these regions are important for theory of mind processing (Saxe & Kanwisher, 2003; Stone, Baron-Cohen, & Knight, 1998; Stuss, Gallup, & Alexander, 2001), and such processing may occur during the reading of dialog. The activation of premotor and sensorimotor cortex for motor imagery clauses supports the hypothesis that readers activate somatomotor representations during the comprehension of motor-related activity.

Speer et al. (2009), in their initial analysis of these data, reported the activation of similar somatomotor regions during clauses in which a character interacted with a new object. The current result for motor imagery clauses clarifies the significance of this previous result: It suggests that when readers update object representations in situation models, they perform motor imagery (Speer et al., 2009). However, it is important to keep in mind that

<sup>3</sup> Fig. S1 presents the brain regions whose activity changed with the onset of low-imagery clauses in Study 1. The reading of low-imagery clauses, compared to baseline, was associated with a decrease in activity in bilateral middle frontal gyrus, bilateral anterior cingulate, and small portions of bilateral precuneus.

the two results are derived from the same data and are not independent. Out of the 148 motor imagery clauses in the current analysis, 60 of them (40.5%) contained object changes.

### 3. Study 2

The results from Study 1 are consistent with the proposal that readers activate modality-specific representations of event features when comprehending discourse. Discourse comprehension theories distinguish between levels of representations that are constructed while reading or listening. A particularly important distinction is between representations of individual clauses or sentences and representations of the larger situation described by a connected discourse (Ferstl et al., 2008; Kintsch, 1998; Van Berkum, Hagoort, & Brown, 1999; Zwaan & Radvansky, 1998). Do readers generally activate sensorimotor regions only when comprehending continuous discourse, or are disconnected sentences sufficient to invoke such activation? There are empirical reasons to believe that disconnected sentences are enough to produce sensorimotor activation. Previous research has shown that readers activate sensorimotor brain regions during the reading of isolated verbs (Hauk et al., 2004; Kable, Lease-Spellmeyer, & Chatterjee, 2002) and the comprehension of short sentences that describe actions (Aziz-Zadeh et al., 2006; Tettamanti et al., 2005). In addition, most of the previous behavioral research that provides evidence for sensorimotor simulation has used single word reading or short sentences (e.g., Glenberg & Kaschak, 2002; Kaschak et al., 2005; Pecher, Zeelenberg, & Barsalou, 2003; Stanfield & Zwaan, 2001; Zwaan & Taylor, 2006; Zwaan et al., 2002). However, in many of these studies the explicit task requirements or implicit task demands may have encouraged a deliberate imagery strategy. As a result, the observed effects could reflect processes engaged by decision requirements specific to the tasks rather than during comprehension itself (Graesser et al., 1997; Mahon & Caramazza, 2008). These limitations temper the conclusions that can be drawn by previous studies regarding the mechanisms of normal comprehension of connected discourse. In Study 1, participants read meaningful connected narratives. An important question is: When the task requirements do not actively promote an imagery strategy, does the text need to form a connected discourse in order for readers to activate modality-specific information? To what extent do people perform sensorimotor simulations when simply reading disconnected sentences for comprehension?

There are theoretical reasons to think why readers *will* activate sensorimotor representations during the comprehension of disconnected sentences, and reasons why they *will not*. On the one hand, a strict adherence to motor-resonance accounts (Zwaan & Taylor, 2006) would suggest that readers will simulate sensorimotor features whenever they are mentioned because the activation of such is critical to understanding any unit of text (Pulvermüller et al., 2005). Any reference to a perceptual feature will activate, or resonate with, the relevant sensorimotor representations. Along these lines, we should observe strong modality-specific effects regardless of whether participants are reading sentences embedded *incoherently* in a larger text (i.e., disconnected sentences), or embedded *coherently* in a larger text so as to form a discourse. On the other hand, it is possible that sensorimotor activation is most likely important when a reader is engaged in generating a globally-coherent situation model of the larger discourse (Fincher-Kiefer, 2001). This is consistent with recent theories suggesting that the activation of sensorimotor representations is most likely to occur in situations when readers are engaged in the construction of meaning (Barsalou, Santos, Simmons, & Wilson, 2008), or are able to develop more elaborate representations of events (Mahon & Caramazza, 2008). Along these lines, we should observe stronger

modality-specific effects when participants read coherent stories than when they read disconnected sentences.

Study 2 asked to what extent readers activate modality-specific representations while reading disconnected sentences, by comparing evoked brain activity during high-imagery clauses during the comprehension of connected discourse to activity during the same clauses presented as isolated sentences. We used an existing fMRI dataset collected by Yarkoni et al. (2008). In the Yarkoni et al. (2008) study, participants read intact stories and scrambled stories in which sentences from different stories were concatenated in random order. Thus, readers were able to construct globally coherent situation models for the intact stories, but not for the scrambled stories. There is evidence that this manipulation is effective in changing a reader's ability to generate a globally-coherent representation of the stories: Yarkoni et al. (2008) found that a region thought important for situation model construction, dorsomedial prefrontal cortex, was more active during the reading of coherent stories than incoherent stories, and comprehension was significantly better for the intact stories. In the current study, we used this dataset to test whether modality-specific representations would be activated only when the stories were presented intact, allowing the reader to construct a coherent discourse representation.

#### 3.1. Method

We will briefly summarize the methods presented in Yarkoni et al. (2008), and will elaborate on the methods specific to the current study.

##### 3.1.1. Neuroimaging methods

Twenty-nine right-handed individuals volunteered to participate in the Yarkoni et al. (2008) study (ages 18–32 years, 17 women). Forty-eight scenes were taken from *One Boy's Day* (Barker & Wright, 1951), the same source as the stories from Study 1, and used to construct sentence sets. For each participant, half of the sentence sets were intact stories (*story* condition) and for the other half sentences from the scenes were sampled quasi-randomly to create scrambled sentence sets (*scrambled* condition). For the story condition, sentence sets had a mean of 132.27 words and 19.31 clauses (*SDs*: 1.83, 1.82). For the scrambled condition, sentence sets had a mean of 132.2 words and 19.31 clauses (*SDs*: 1.95, 1.68). The assignment of sentence to the story or scrambled condition was counterbalanced across participants with two lists. Each list contained 24 intact story sentence sets and 24 scrambled sentence sets (see Yarkoni et al., 2008, for more details). The Yarkoni et al. (2008) stimuli can be downloaded from <http://dcl.wustl.edu/DCL/stimuli.html>.

Narratives were presented one word at a time. Each word was presented for 200 ms with a 100 ms ISI. An additional 400 ms delay followed the end of a sentence, resulting in a 500 ms ISI between sentences. The stories were presented in 12 runs. Each run consisted of the presentation of four blocks of sentences, two of which were intact stories and two of which were scrambled sentences. Before each block, participants were presented with an instruction screen, for 4 s, indicating whether they were going to be reading intact or scrambled stories. For the scrambled condition, they were explicitly told that there was no reason to try to construct a coherent story from the sentences. The order of blocks was counterbalanced across participants. Participants were instructed to read and understand the sentences for a later memory test.

fMRI data acquisition and preprocessing was similar to Study 1. Images were acquired on a 3T Siemens Vision MRI scanner (Erlangen, Germany). Structural images were acquired using a sagittal MP-RAGE T1-weighted sequence (1 × 1 × 1.25 mm) and a T2-weighted fast turbo spin-echo scan, and functional images were

acquired using a  $T2^*$ -weighted fast turbo spin-echo echo-planar sequence, with 32 slices ( $4.0 \times 4.0$  mm in-plane resolution), acquired every 2.048 s. The functional data were preprocessed following standard protocols (see Yarkoni et al. (2008) for details) and warped to a standard stereotactic space with 3 mm isotropic voxels (Talairach & Tournoux, 1988).

### 3.1.2. Collection of imagery norms and imagery modality coding

We collected imagery norms and coded imagery modality using the same procedures as were used in Study 1. We used the same story and scrambled sentence sets as in Yarkoni et al. (2008). In the norming study, a new set of 32 participants read each story one clause at a time, and rated their experienced strength of imagery after each clause, using the same 1 (no imagery at all) to 6 (as vivid as perception) scale as in Study 1. As in Yarkoni et al. (2008), sentences were presented in 4-text runs: two story blocks and two scrambled blocks. Participants were informed as to whether the upcoming block was a set of intact stories or scrambled stories. The presentation of story order and assignment of sentence to condition was counterbalanced as in Yarkoni et al. (2008).

To code for imagery modality, we computed the mean imagery strength for each clause by coherence condition (intact story or scrambled) and counterbalance list.<sup>4</sup> We then selected the clauses, per story type and counterbalance list, in the top two quintiles of imagery strength for modality coding. (There were a total of 927 clauses per counterbalance list, with a subset of those selected for modality coding.) A small percentage (11.1%) of the clauses were identical to those used in Study 1. We used the modality codes from Study 1 for those clauses if they were selected as high-imagery clauses in the new imagery norms. For the coherent story condition, there were 56 visual imagery, 84 motor imagery, and 42 auditory imagery clauses for counterbalance list 1, and 64 visual imagery, 94 motor imagery, and 43 auditory imagery clauses for counterbalance list 2. For the incoherent story condition, there were 59 visual imagery, 95 motor imagery, and 38 auditory imagery clauses for counterbalance list 1, and 59 visual imagery, 105 motor imagery, and 31 auditory imagery clauses for counterbalance list 2.<sup>5</sup> The remaining clauses (558 for list 1 and 531 for list 2) were categorized as low-imagery, as in Study 1.

### 3.1.3. Analysis of the neuroimaging data

The data analysis procedure was similar to that of Study 1. Four predictor variables coded the clause onset of each of the three imagery types (auditory, motor, and visual imagery) and the low-imagery control for the story condition, and four predictor variables coded for the clause onset for the three imagery types and the low-imagery control for the scrambled condition. Two regressors were created to code for the onset of the instruction screens for the story and scrambled conditions. To generate regressors for the GLMs, all eight clause types were each coded as a brief pulse of activity at the beginning of the clause, and the two instruction screen regressors were coded with a 4 s duration, all of which were convolved with a canonical hemodynamic response function (Boynton et al., 1996). Ten additional regressors were included to code for effects of no interest: terminal and non-terminal punctuation, differences between runs, and linear trend within each run.

<sup>4</sup> A preliminary inspection of the ratings indicated that the mean imagery strength was different between the two counterbalance lists, so we computed means separately within each list.

<sup>5</sup> Similar to Study 1, there were not enough instances of haptic, olfactory, and gustatory imagery to warrant their inclusion in the analyses (Story: 3 haptic, 0 olfactory, 3 gustatory for counterbalance list 1, and 0 haptic, 1 olfactory, 0 gustatory for counterbalance list 2; Scrambled: 0 haptic, 1 olfactory, 0 gustatory for counterbalance list 1, and 3 haptic, 0 olfactory, 2 gustatory for counterbalance list 2).

### 3.1.4. Region-of-interest analyses

We first conducted region-of-interest analyses using the regions that showed significant imagery effects in Study 1 (presented in Table 1) to test whether the imagery effects from Study 1 replicated in this new dataset, and whether their responses differed between the story and scrambled conditions. For each modality, we used the corresponding regions that showed a significant imagery effect in Study 1 (e.g., auditory imagery regions for the auditory imagery analyses, etc.). We conducted paired-sample *t* tests, with subject as a random factor, comparing the effect estimates for each of the three coherent story modality types with the coherent story low-imagery variable ( $\text{Auditory}_{\text{story}} - \text{LowImagery}_{\text{story}}$ ,  $\text{Motor}_{\text{story}} - \text{LowImagery}_{\text{story}}$ , and  $\text{Visual}_{\text{story}} - \text{LowImagery}_{\text{story}}$ ) and compared effect estimates for each of the three scrambled story modality types with the scrambled story low-imagery variable ( $\text{Auditory}_{\text{scrambled}} - \text{LowImagery}_{\text{scrambled}}$ ,  $\text{Motor}_{\text{scrambled}} - \text{LowImagery}_{\text{scrambled}}$ , and  $\text{Visual}_{\text{scrambled}} - \text{LowImagery}_{\text{scrambled}}$ ). To test for an imagery by coherence interaction – whether the strength of the imagery effects were stronger for the story condition than the scrambled condition – we computed paired-samples *t* tests comparing the difference between each story imagery and story control contrast with the difference between the corresponding scrambled imagery and scrambled control contrast (e.g.,  $([\text{Auditory}_{\text{story}} - \text{LowImagery}_{\text{story}}] - [\text{Auditory}_{\text{scrambled}} - \text{LowImagery}_{\text{scrambled}}])$ , and so on). For all analyses, significance levels were Bonferroni corrected using the number of comparisons within each modality type.

### 3.1.5. Whole-brain analyses

We tested the same sets of contrasts as were used in the region-of-interest analyses in a whole-brain analysis. To find regions that responded to the different imagery modalities regardless of coherence condition, we computed a contrast for each modality that averaged over the story and scrambled conditions (e.g.,  $\text{Auditory}_{\text{Average}} - \text{LowImagery}_{\text{Average}}$ , and so on). Separate contrasts for scrambled and story conditions, as well as interactions between imagery strength and coherence were tested within these regions. Maps of *t*-statistics were generated and converted to *z*-statistic maps, which were thresholded to control for map-wise false positive rate at .05 by retaining clusters of at least 2 voxels with *z* values greater than 4.5 (McAvoy et al., 2001). For interaction contrasts, significance levels were Bonferroni corrected using the number of comparisons within each modality type.

## 3.2. Results and discussion

### 3.2.1. Imagery ratings

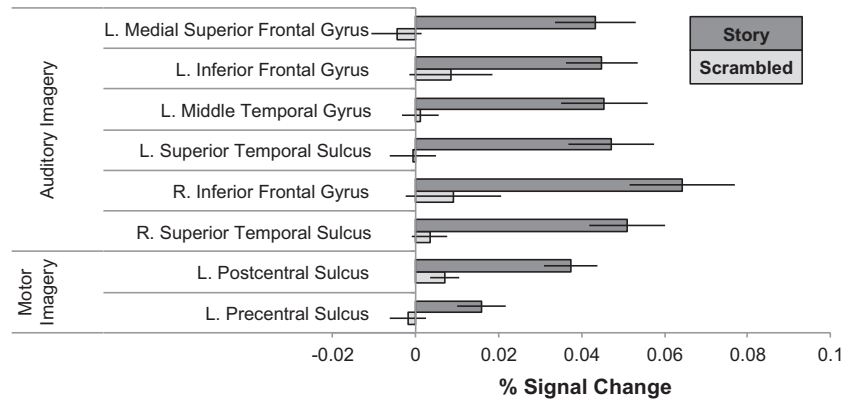
A paired-samples *t* test, with subject as a random factor, indicated that imagery ratings were significantly higher for the story condition ( $M = 3.72$ ,  $SD = 0.85$ ) than the scrambled condition ( $M = 3.15$ ,  $SD = 0.75$ ),  $t(31) = 7.64$ ,  $p < .001$ ,  $d = 1.35$ . This shows that the experienced strength of imagery was higher when reading intact than scrambled stories.

### 3.2.2. fMRI results

We first present the region-of-interest analysis results, and then the whole-brain results.

### 3.2.3. Region-of-interest analysis results

Table 1 presents the results of the analyses for each coherence condition and for the coherence by imagery interaction for each region of interest from Study 1. For all imagery modalities, none of the regions from Study 1 showed significant imagery effects in the scrambled story condition. However, many of these regions showed significant imagery effects for the story condition.



**Fig. 3.** Mean signal change for regions from Study 1 that had a significant coherence by imagery interaction in Study 2. All regions showed greater activity for the story than the scrambled condition. The error bars indicate  $\pm 1$  standard error.

For the reading of auditory imagery clauses, all but two regions from Study 1 showed significant imagery effects in the story condition. Additionally, a majority of those regions showed an interaction between imagery strength and coherence such that they showed a larger increase in activity in the story condition than in the scrambled story condition (see Fig. 3).

For motor imagery clauses, both sensorimotor regions showed significant imagery effects in the story condition, and showed a significantly larger effect for the story condition than the scrambled story condition (see Fig. 3).

These results replicate the results from Study 1. In addition, across the three modalities most regions showed an imagery by coherence interaction such that imagery effects were stronger for the story than scrambled condition. This suggests that readers activate modality-specific representations more strongly during discourse comprehension when they are able to develop discourse-level story representations.

### 3.2.4. Whole-brain analysis results

The regions that showed imagery effects, collapsed across coherence condition, are illustrated in Fig. 4 and presented in Table 2.<sup>6</sup> Table 2 also shows whether each region showed a significant effect separately for the scrambled and story conditions, and whether that region showed a significant imagery strength  $\times$  coherence interaction. Most regions showed significant imagery effects for the story condition, and only two regions showed effects in the scrambled condition. There was a large amount of overlap between the regions showing significant effects in this whole-brain analysis and those regions showing imagery effects in Study 1.

For auditory imagery clauses, there were significant imagery effects in large portions of the temporal lobe bilaterally. This activation includes regions in secondary auditory cortex (BA 22) such as the left posterior superior temporal gyrus at the temporo-parietal junction and superior temporal sulcus. Similar to Study 1, there were no effects in the right TPJ. Additionally, similar to Study 1, there were significant effects in bilateral inferior frontal cortex. All except for four of the regions listed in Table 2 showed an imagery strength by coherence interaction such that effects were stronger for the story than scrambled story condition.

<sup>6</sup> Fig. S2 presents the brain regions whose activity changed with the onset of low-imagery clauses in Study 2. The reading of low-imagery clauses, compared to baseline, was associated with a decrease in activity in bilateral middle frontal and superior frontal gyrus, bilateral insular cortex, bilateral inferior parietal lobule, bilateral precuneus and cuneus, bilateral anterior cingulate gyrus, and bilateral medial frontal cortex. Compared to baseline, the reading of low-imagery clauses was also associated with increases in activity in bilateral lateral occipital cortex, bilateral middle temporal gyrus, bilateral precentral gyrus, and left inferior frontal gyrus.

For the reading of motor imagery clauses, there was a significant imagery effect in a portion of the left postcentral sulcus. This region showed a larger increase for the story condition than the scrambled story condition.

As in Experiment 1, there were no significant effects for visual imagery clauses.

In sum, the region-of-interest and whole-brain analyses converge to show a replication of the auditory-specific and motor-specific imagery effects found in Study 1. Readers tended to activate sensorimotor regions relevant for the imagery information implied in the text. Additionally, the data suggest that imagery strength depends on discourse-level processing. Imagery ratings were higher for the story condition than the scrambled condition. For the fMRI analyses, there were very few regions that had significant imagery effects in the scrambled story condition, and most regions showed significantly stronger imagery effects for the story condition than the scrambled condition.

## 4. General discussion

In these two studies, we tested whether the activation of modality-specific representations occurs during extended discourse comprehension. We found evidence supporting this possibility: In both studies, readers activated sensorimotor regions relevant to the perceptual information implied in the text. In addition, the behavioral and imaging data from Study 2 showed that the modality-specific effects only occurred when participants were able to develop discourse-level representations. In both studies, readers' only goal was to read for comprehension and memory; they did not engage in explicit judgment tasks during reading. The results thus provide good evidence that the activation of modality-specific representations occurs during the ongoing reading of discourse and is not merely the product of task demands.

Across both studies, reading clauses with high auditory imagery content was associated with increases in regions within secondary auditory cortex, including left superior temporal gyrus and bilateral superior temporal sulcus. Additionally, we found activation of perisylvian language-related regions associated with the reading of auditory imagery clauses. Specifically, we found activation of left posterior superior temporal gyrus at the TPJ and left inferior frontal gyrus, which correspond roughly to the classic Wernicke's and Broca's areas (Catani et al., 2005; Ferstl et al., 2008). (The activation of inferior frontal gyrus was bilateral.) Overall, these effects are consistent with previous research on the access of sound knowledge and auditory imagery (Halpern & Zatorre, 1999; Halpern, Zatorre, Bouffard, & Johnson, 2004; Hubbard, 2010; McGuire et al., 1996;



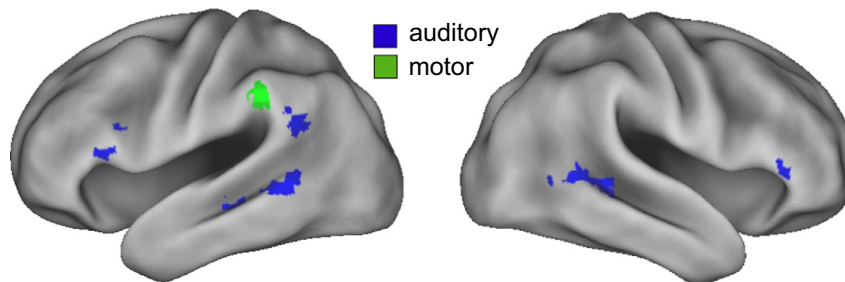


Fig. 4. Regions showing modality-specific imagery effects for Study 2, collapsed across coherence condition.

Table 2

Regions from the Study 2 whole-brain analysis that increased in activity for the different imagery modalities.

Imagery modality	Center of mass coordinates			Hemisphere	Region	Volume (mm <sup>3</sup> )	Peak z statistic	Study 2 scrambled	Study 2 story	Study 2 imagery X coherence interaction
	x	y	z							
Auditory	–57	–51	27	Left	Posterior superior temporal gyrus, at temporo-parietal junction (BA 22/40)	18	4.67	–	✓	–
	–53	–54	21	Left	Superior temporal gyrus (BA 22)	126	5.16	–	✓	✓
	–54	–43	–3	Left	Middle temporal gyrus (BA 22)	792	5.99	✓	✓	✓
	–54	20	10	Left	Inferior frontal gyrus (BA 45)	207	5.77	–	✓	✓
	–45	11	21	Left	Middle frontal gyrus (BA 44/45)	27	4.67	–	✓	–
	–12	40	54	Left	Superior frontal gyrus (BA 8)	54	5.31	✓	✓	–
	–11	19	62	Left	Superior frontal gyrus (BA 6)	81	5.38	–	✓	–
	51	28	6	Right	Inferior frontal gyrus (BA 47)	81	5.19	–	✓	✓
	52	–44	–1	Right	Superior temporal sulcus (BA 22)	738	5.65	–	✓	✓
Motor	–62	–37	35	Left	Postcentral sulcus (BA 5/40)	621	6.21	✓	✓	✓
Visual	–	–	–	–	–	–	–	–	–	–

Note: Coordinates listed are the center of mass coordinates in Talairach space. Check marks indicate whether the region showed a significant effect for the scrambled condition, story condition, and a imagery X condition interaction, corrected for multiple comparisons using the Bonferroni method (9 regions for auditory imagery).

Shergill et al., 2001; Yoo, Lee, & Choi, 2001; Zatorre, Halpern, Perry, Meyer, & Evans, 1996). They suggest that readers activate auditory representations when reading about sounds and dialog in discourse.

We found activation of secondary somatosensory and premotor cortex associated with the reading of clauses that imply motor information. These results are consistent with a large body of research showing the activation of such regions during motor imagery and observation (Ehrsson, Geyer, & Naito, 2003; Grèzes & Decety, 2001; Michelon, Vettel, & Zacks, 2006; Porro et al., 1996). These results suggest that readers were simulating how to execute the actions described in the text and how it feels to perform them. In addition, these simulations may have been dominant-hand specific; all of the participants were right-handed. Recent research supports this possibility: Willems, Hagoort, et al. (2010) and Willems, Toni, et al. (2010) found that when reading hand-action verbs, left-handers preferentially activated right motor cortex whereas right-handers preferentially activated left motor cortex.

We did not find effects of reading visual imagery clauses for either study. One possibility is that the concurrent visual demands of reading reduced the sensitivity of these studies to detect visual imagery effects. It would be informative for future studies to compare the magnitude of visual imagery effects during reading to visual imagery effects during listening.

An important point about these interpretations is that they depend in part on “reverse inference”—going from observation of the location of an observed brain effect to a conclusion about the sort of processing that is being performed. Reverse inference has been rightly criticized as being inferentially weak in many cases (Van Horn & Poldrack, 2009). However, the strength of reverse inference

depends on the selectivity of activation in the brain area in question; an area that (a) reliably activates for a particular type of processing, and (b) does not reliably activate for other types of processing does license strong reverse inference (Poldrack, 2006). In the present case, the activity in perisylvian language areas and the simultaneous activation in premotor and parietal areas corresponding to hand representations is moderately selective, and thus we are moderately confident in inferring function from these activations. Our confidence in the validity of the functional inference is strengthened by the fact that regions identified based on Study 1 replicated in an independent data set in Study 2, and by the fact that these regions and others in the whole-brain analysis were modulated by text coherence.

It has been argued that readers generate sensorimotor simulations in part to construct situation models, which are globally coherent representations of the events described by the story (Zwaan, 2004; Zwaan & Radvansky, 1998). This suggests that sensorimotor activation should be stronger when readers are able to generate situation models of the narrative than when they are not able to do so. The results from Study 2 support this possibility; most of the regions that showed a main effect of imagery also exhibited a coherence by imagery interaction showing that the fMRI sensorimotor activation effects were significantly stronger in the story than scrambled condition. These results suggest that situation models of narrative discourse have perceptual contents. Some behavioral research supports this possibility. Fincher-Kiefer (2001), using Albrecht & O'Brien, 1993 contradiction paradigm, had participants read narratives that ended in either a consistent or contradictory way than implied by the story. For example, participants read a story about Mary entering a restaurant for lunch. In

one version of the story, Mary is described as a fast food junky, and in another version she is described as being a vegetarian. If Mary then acts in a contradictory way, e.g., ordering a cheeseburger and fries when she was described as a vegetarian, reading time slows down in comparison to when she orders in an expected way. Fincher-Kiefer (2001) found that readers did not show these contradiction effects when concurrently maintaining high-imagery sentences but did when maintaining low-imagery sentences. This supports the possibility that global coherence building draws upon imagery processes. However, whether discourse comprehension relies on imagery or merely co-occurs with it is currently unknown and future research is needed to assess this question.

Although the current study found little evidence for activation of sensorimotor representations during the comprehension of disconnected sentences, this is in apparent contrast with previous studies finding imagery effects using isolated words or short texts. There are two important differences between our studies and those previous ones. First, the majority of the previous studies finding imagery effects using isolated units of text had participants engage in explicit judgment tasks of the language materials, which may promote imagery (but see, for example, Tettamanti et al., 2005). The presence of such tasks makes it difficult to determine if readers would *naturally* generate images for such materials. Second, the critical factor in the activation of sensorimotor representations in language comprehension may be the ability to create coherent mental models, regardless of text length. The scrambled condition in Study 2 used texts that should be particularly difficult to organize into coherent mental models; unrelated sentences arranged randomly into paragraphs. For example, the sentence “Raymond greeted her in a friendly voice, ‘Hi, Honey.’” might not afford sensorimotor simulation when taken out of its discourse context, in which it was normally preceded by, “Honey, Raymond’s fat, broad, elderly fox terrier, ambled into the room.” In at least some previous studies, isolated sentences may have afforded the construction of a coherent model, for example, “Raymond greeted his dog in a friendly voice, ‘Hi, Honey.’” Such an isolated sentence, although short, presents a coherent event and supports the construction of a coherent mental model of that event. Major theories of text comprehension argue that readers are driven by a need to build coherence (Graesser, Singer, & Trabasso, 1994), and will attempt to do so even if the textual unit is very small (e.g., a single word or short phrase). As such, readers will build coherent mental models of very short textual input. If larger sets of text are presented, such as in discourse, readers will attempt to build coherence from the larger input. If readers are prevented from building coherence from the discourse, as in our present study, coherence processes will likely be disrupted as well as, we argue, the sensorimotor processing that may occur with them.

Do sensorimotor representations contribute causally to comprehension or do they merely co-occur with it? This question still remains largely unanswered (Mahon & Caramazza, 2008) though some have made arguments that the activation of sensorimotor representations is critical to understanding (Glenberg & Gallese, 2012). Our data cannot say much beyond that such representations co-occur with comprehension. However, our data suggest that the activation of sensorimotor representations is not obligatory during sentence processing. It depends on whether participants can generate coherent situation models. This suggests that situation models contain perceptual contents, which is consistent with theoretical accounts of how perceptual simulations may contribute to comprehension: they make event understanding richer (Mahon & Caramazza, 2008).

It is tempting to view the results from the current set of studies as leverage in the debate regarding whether knowledge is inherently perceptual or amodal. Although the activation of sensorimotor systems during comprehension is certainly consistent with the

proposal that readers activate sensorimotor simulations, it does not rule out the possibility that readers also activate abstract content-specific knowledge (Louwerse, 2008; Mahon & Caramazza, 2008). Mahon and Caramazza (2008) argue for the possibility that simulations are due to activation cascading from higher-level abstract representations. In order to test this possibility, new studies would need to assess the relative timing of the activation of sensorimotor systems and non-perceptual systems, which could be tested with higher temporal resolution systems.

Does modality-specific activity during comprehension mean that readers are generating simulations – reenactments of sensorimotor states – or activating content specific representations containing sensorimotor information? A strict interpretation of sensorimotor simulation theories suggests that a simulation occurs when readers activate sensorimotor systems in the *right way*. That is, the activation of sensorimotor systems when reading about actions, visual information, sounds, etc. should be topographically organized. For example, a simulation of the action “I grasp the ball” would be best characterized by the neural computations important to executing the fine motor actions involved in grasping. The current study showed activity outside of somatotopically-organized regions only. Our current results, then, support a weaker version of a simulation hypothesis: that readers activate modality-specific representations during comprehension. This weaker version is consistent with a recent study showing that although both explicit imagery of verbs and the reading of verbs evoked somatotopically organized activity, the regions did not overlap across tasks (Willem, Toni, et al., 2010). Our findings are also in contrast to theories positing automatic activation of motor cortices when reading about actions (Pulvermüller, 1999; Pulvermüller et al., 2005). Of great importance for future studies is to systematically test whether or not sensorimotor activity when reading about perceptual information is topographically organized, using localizers and experimental manipulation of fine perceptual detail in narrative events.

The reading task used in this study is considerably more naturalistic than in most previous research. However, in order to measure brain activity with fMRI the typical reading situation had to be adjusted. Most important, participants read the texts one word at a time while lying in a scanner. Free viewing, while more characteristic of normal reading, introduces difficult confounds due to eye movements. To minimize the disruption of normal reading processes, participants practiced this style of reading before the scanning session and reported feeling comfortable with the task. That these reading conditions may be less than natural makes the results quite revealing. Results under these conditions may well underestimate the amount of sensorimotor representation readers engage during fully natural reading experiences.

Narratives typically describe events with rich perceptual detail. Readers may generate images of such detail during comprehension, and may rely on modality-specific systems to do so. When reading about visual scenes, they may generate visual images about those scenes. When reading about sounds, they may generate auditory images of those sounds. When reading about actions, they may generate somatomotor images about what it is like to perform those actions. The present results suggest that such sensorimotor representations may participate meaningfully in comprehension: They appear to occur during ongoing discourse comprehension, without explicit judgment tasks, when readers are able to generate globally coherent situation models.

## Acknowledgments

Preparation of this manuscript was partially supported by Grants T32 AG000030-31 and RO1-MH70674 from the National Institutes of Health.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bandl.2013.07.003>.

## References

- Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning*, 19, 1061–1070.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, 4, 267–278.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Aziz-Zadeh, L., Wilson, S., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, 16, 1818–1823.
- Barker, R. G., & Wright, H. S. (1951). *One boy's day; a specimen record of behavior*. Oxford, England: Harper.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–660.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.
- Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 245–283). Oxford, England: Oxford University Press.
- Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., & Saxe, R. (2008). Concepts are more than percepts: The case of action verbs. *The Journal of Neuroscience*, 28, 11347–11353.
- Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience*, 16(13), 4207–4221.
- Brunyé, T. T., Ditman, T., Mahoney, C. R., Walters, E. K., & Taylor, H. A. (2010). You heard it here first: Readers mentally simulate described sounds. *Acta Psychologica*, 135(2), 209–215. doi: <http://dx.doi.org.ezproxy.gvsu.edu/10.1016/j.actpsy.2010.06.008>.
- Castiello, U. (2005). The neuroscience of grasping. *Nature Reviews Neuroscience*, 6, 726–736.
- Catani, M., Jones, D. K., & ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology*, 57, 8–16.
- Damasio, H., Grabowski, T. J., Tranel, D., Pronto, L. L., Hichwa, R. D., & Damasio, A. R. (2001). Neural correlates of naming actions and of naming spatial relations. *Neuroimage*, 13, 1053–1064.
- Desai, R. H., Binder, J. R., Conant, L. L., & Seidenberg, M. S. (2010). Activation of sensory-motor areas in sentence comprehension. *Cerebral Cortex*, 20, 468–478.
- D'Esposito, M., Detre, J. A., Aguirre, G. K., Stallcup, D., Alsop, D. C., Tippet, L. J., et al. (1997). A functional MRI study of mental image generation. *Neuropsychologia*, 35, 725–730.
- 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.
- Ferstl, E. C., Neumann, J., Bogler, C., & von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping*, 29, 581–593.
- Fincher-Kiefer, R. (2001). Perceptual components of situation models. *Memory & Cognition*, 29, 336–343.
- Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of motor system in language comprehension. *The Quarterly Journal of Experimental Psychology*, 61(6), 825–850. doi: <http://dx.doi.org.ezproxy.gvsu.edu/10.1080/17470210701623605>.
- Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
- Fries, U., Rutschmann, R., Raabe, M., & Schmalhofer, F. (2008). Neural indicators of inference processes in text comprehension: An event-related functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, 20(11), 2110–2124. doi: <http://dx.doi.org.ezproxy.gvsu.edu/10.1162/jocn.2008.20141>.
- Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22(3–4), 455–479. doi: <http://dx.doi.org.ezproxy.gvsu.edu/10.1080/02643290442000310>.
- Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, 20, 1–55.
- Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 48(7), 905–922. doi: <http://dx.doi.org.ezproxy.gvsu.edu/10.1016/j.cortex.2011.04.010>.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9, 558–565.
- Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, 48, 163–189.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371–395. doi: <http://dx.doi.org.ezproxy.gvsu.edu/10.1037/0033-295X.101.3.371>.
- Grèzes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping*, 12, 1–19.
- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., et al. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, 12, 711–720.
- Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, 9, 697–704.
- Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, 42, 1281–1292.
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41, 301–307.
- Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, 6, 242–247.
- Hubbard, T. L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin*, 136, 302–329.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Kable, J. W., & Chatterjee, A. (2006). Specificity of action representations in the lateral occipitotemporal cortex. *Journal of Cognitive Neuroscience*, 18, 1498–1517.
- Kable, J. W., Lease-Spellmeyer, J., & Chatterjee, A. (2002). Neural substrates of action event knowledge. *Journal of Cognitive Neuroscience*, 14, 795–805.
- Kanwisher, N., Downing, P., Epstein, R., & Kourtzi, Z. (2001). Functional neuroimaging of human visual recognition. In R. Cabeza & A. Kingstone (Eds.), *The handbook of functional neuroimaging* (pp. 109–152). Cambridge, MA: MIT Press.
- Kaschak, M. P., Madden, C. J., Theriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., et al. (2005). Perception of motion affects language processing. *Cognition*, 94, B79–B89.
- Kellenbach, M. L., Brett, M., & Patterson, K. (2001). Large, colorful, or noisy? Attribute- and modality-specific activations during retrieval of perceptual attribute knowledge. *Cognitive, Affective & Behavioral Neuroscience*, 1, 207–221.
- Kemmerer, D., & Gonzalez-Castillo, J. (2010). The two-level theory of verb meaning: An approach to integrating the semantics of action with the mirror neuron system. *Brain and Language*, 112(1), 54–76. doi: <http://dx.doi.org.ezproxy.gvsu.edu/10.1016/j.bandl.2008.09.010>.
- Kiefer, M., Sim, E.-J., Herrnberger, B., Grothe, J., & Hoenig, K. (2008). The sound of concepts: Four markers for a link between auditory and conceptual brain systems. *The Journal of Neuroscience*, 28(47), 12224–12230. doi: <http://dx.doi.org.ezproxy.gvsu.edu/10.1523/JNEUROSCI.3579-08.2008>.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY, US: Cambridge University Press.
- Kurby, C. A., Magliano, J. P., & Rapp, D. N. (2009). Those voices in your head: Activation of auditory images during reading. *Cognition*, 112, 457–461.
- Louwerse, M. M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin & Review*, 15, 838–844.
- Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition*, 114, 96–104.
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology – Paris*, 102, 59–70.
- Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, 3, 173–192.
- Markman, A. B., & Brendl, C. M. (2005). Constraining theories of embodied cognition. *Psychological Science*, 16, 6–10.
- Marks, D. F. (1995). New directions for mental imagery research. *Journal of Mental Imagery*, 19(3–4), 153–167.
- McAvoy, M., Ollinger, J. M., & Buckner, R. L. (2001). Cluster size thresholds for assessment of significant activation in fMRI. *Neuroimage*, 13, S198.
- McGuire, P. K., Silbersweig, D. A., Murray, R. M., David, A. S., Frackowiak, R. S. J., & Frith, C. D. (1996). Functional neuroanatomy of inner speech and auditory verbal imagery. *Psychological Medicine*, 26, 29–38.
- Michelon, P., Vettel, J. M., & Zacks, J. M. (2006). Lateral somatotopic organization during imagined and prepared movements. *Journal of Neurophysiology*, 95, 811–822.
- Orban, G. A., Van Essen, D., & Vanduffel, W. (2004). Comparative mapping of higher visual areas in monkeys and humans. *Trends in Cognitive Sciences*, 8, 315–324.
- Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science*, 14, 119–124.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10, 59–63.
- Porro, C. A., Francescato, M. P., Cettolo, V., Diamond, M. E., et al. (1996). Primary motor and sensory cortex activation during motor performance and motor imagery: A functional magnetic resonance imaging study. *The Journal of Neuroscience*, 16, 7688–7698.
- Postle, N., McMahon, K. L., Ashton, R., Meredith, M., & de Zubicaray, G. I. (2008). Action word meaning representations in cytoarchitecturally defined primary and premotor cortices. *NeuroImage*, 43, 634–644.
- Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, 22(2), 253–336. doi: <http://dx.doi.org.ezproxy.gvsu.edu/10.1017/S0140525X9900182X>.

- Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, 21, 793–797.
- Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, 88, 16–45.
- Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, 25, 157–238.
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in “theory of mind”. *NeuroImage*, 19(4). [http://dx.doi.org/10.1016/S1053-8119\(03\)00230-1](http://dx.doi.org/10.1016/S1053-8119(03)00230-1).
- Shergill, S. S., Bullmore, E. T., Brammer, M. J., Williams, S. C. R., Murray, R. M., & McGuire, P. K. (2001). A functional study of auditory verbal imagery. *Psychological Medicine*, 31, 241–253.
- Simmons, W. K., Ramjee, V., Beauchamp, M. S., McArae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia*, 45, 2802–2810.
- Speer, N. K., Reynolds, J. R., Swallow, K. M., & Zacks, J. M. (2009). Reading stories activates neural representations of visual and motor experiences. *Psychological Science*, 20, 989–999.
- Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological Science*, 18, 449–455.
- Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12, 153–156.
- Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience*, 10(5). <http://dx.doi.org/10.1162/089892998562942>.
- Stuss, D. T., Gallup, G. G. J., & Alexander, M. P. (2001). The frontal lobes are necessary for “theory of mind”. *Brain: A Journal of Neurology*, 124(Pt 2). <http://dx.doi.org/10.1093/brain/124.2.279>.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system, an approach to cerebral imaging*. Stuttgart: G. Thieme.
- 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.
- Thompson-Schill, S. L., Aguirre, G. K., D’Esposito, M., & Farah, M. J. (1999). A neural basis for category and modality specificity of semantic knowledge. *Neuropsychologia*, 37, 671–676.
- Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, 11, 657–671.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies in discourse comprehension*. New York, NY, US: Academic Press.
- Van Essen, D. C. (2005). A population-average, landmark- and surface-based (PALS) atlas of human cerebral cortex. *NeuroImage*, 28, 635–662.
- Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of American Medical Informatics Association*, 8, 443–459.
- Van Horn, J. D., & Poldrack, R. A. (2009). Functional MRI at the crossroads. *International Journal of Psychophysiology*, 73, 3–9.
- Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory’s echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences*, 97, 11125–11129.
- Willems, R. M., Hagoort, P., & Casasanto, D. (2010). Body-specific representations of action verbs: Neural evidence from right- and left-handers. *Psychological Science*, 21, 67–74.
- Willems, R. M., Labruna, L., D’Esposito, M., Ivry, R., & Casasanto, D. (2011). A functional role for the motor system in language understanding: Evidence from theta-burst transcranial magnetic stimulation. *Psychological Science*, 22, 849–854.
- Willems, R. M., Toni, I., Hagoort, P., & Casasanto, D. (2010). Neural dissociations between action verb understanding and motor imagery. *Journal of Cognitive Neuroscience*, 22(10). <http://dx.doi.org/10.1162/jocn.2009.21386>.
- Yao, B., Belin, P., & Scheepers, C. (2011). Silent reading of direct versus indirect speech activates voice-selective areas in the auditory cortex. *Journal of Cognitive Neuroscience*, 23, 3146–3152. doi: [http://dx.doi.org.ezproxy.gvsu.edu/10.1162/jocn\\_a.00022](http://dx.doi.org.ezproxy.gvsu.edu/10.1162/jocn_a.00022).
- Yarkoni, T., Speer, N. K., & Zacks, J. M. (2008). Neural substrates of narrative comprehension and memory. *NeuroImage*, 41, 1408–1425.
- Yoo, S.-S., Lee, C. U., & Choi, B. G. (2001). Human brain mapping of auditory imagery: Event-related functional MRI study. *NeuroReport*, 12, 3045–3049.
- Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., et al. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, 4, 651–655.
- Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., & Evans, A. C. (1996). Hearing in the mind’s ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience*, 8, 29–46.
- Zwaan, R. A. (2004). The immersed experienter: Toward an embodied theory of language comprehension. In Ross, Brian H. (Ed.), *The psychology of learning and motivation: Advances in research and theory* (vol. 44, Academic Press, New York, pp. 35–62).
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162–185.
- Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science*, 28, 611–619.
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shape of objects. *Psychological Science*, 13, 168–171.
- Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, 135, 1–11.
- Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review*, 10, 954–958.