



# Top-down and bottom-up contributions to understanding sentences describing objects in motion

Shirley-Ann Rueschemeyer<sup>1,2\*</sup>, Arthur M. Glenberg<sup>3</sup>, Michael P. Kaschak<sup>4</sup>, Karsten Mueller<sup>1</sup> and Angela D. Friederici<sup>1,5</sup>

<sup>1</sup> Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

<sup>2</sup> Donders Centre for Cognition, Radboud University, Nijmegen, Netherlands

<sup>3</sup> Laboratory for Embodied Cognition, School of Biological and Health Systems Engineering, Arizona State University, School of Life Sciences, Tempe, AZ, USA

<sup>4</sup> Departments of Neuroscience and Communicative Disorders, Florida State University, Gainesville, FL, USA

<sup>5</sup> Center for Advanced Study in the Behavioral Sciences at Stanford University, Stanford, CA, USA

## Edited by:

Diane Pecher, Erasmus University  
Rotterdam, Netherlands

## Reviewed by:

Michael E. J. Masson, University of  
Victoria, Canada

Christine Wilson, Emory University,  
USA

## \*Correspondence:

Shirley-Ann Rueschemeyer, Donders  
Centre for Cognition, Radboud  
University, Montessorilaan 3, 6525 HR,  
Nijmegen, Netherlands.  
e-mail: s.rueschemeyer@donders.ru.nl

Theories of embodied language comprehension propose that the neural systems used for perception, action, and emotion are also engaged during language comprehension. Consistent with these theories, behavioral studies have shown that the comprehension of language that describes motion is affected by simultaneously perceiving a moving stimulus (Kaschak et al., 2005). In two neuroimaging studies, we investigate whether comprehension of sentences describing moving objects activates brain areas known to support the visual perception of moving objects (i.e., area MT/V5). Our data indicate that MT/V5 is indeed selectively engaged by sentences describing objects in motion toward the comprehender compared to sentences describing visual scenes without motion. Moreover, these sentences activate areas along the cortical midline of the brain, known to be engaged when participants process self-referential information. The current data thus suggest that sentences describing situations with potential relevance to one's own actions activate both higher-order visual cortex as well brain areas involved in processing information about the self. The data have consequences for embodied theories of language comprehension: first, they show that perceptual brain areas support sentential-semantic processing. Second the data indicate that sensory-motor simulation of events described through language are susceptible to top-down modulation of factors such as relevance of the described situation to the self.

**Keywords: embodiment, sentence comprehension, self-referentiality, visual motion**

According to embodiment theory, language comprehension requires neural resources ordinarily used for perception, action, and emotion. Thus, for example, understanding language about actions (e.g., "Open the drawer.") requires simulation using neural systems involved in action (i.e., ventral premotor and inferior parietal cortex), whereas comprehension of language about visual motion (e.g., "The car approached you.") requires access to neural systems involved in motion perception (i.e., posterior lateral temporal cortex). The predictions regarding action-related language have been confirmed behaviorally (e.g., Glenberg and Kaschak, 2002), using imaging techniques (e.g., Hauk et al., 2004; Rueschemeyer et al., 2007), and using TMS (Buccino et al., 2005; Glenberg et al., 2008). The case for language denoting moving objects appears less clear cut.

Behavioral evidence has been reported supporting a link between neural systems involved in perception of visual motion and comprehension of language describing objects in motion. For example, Kaschak et al. (2005) showed that participants watching a video clip depicting motion are slower to respond to an acoustically presented sentence describing an object in motion, when the trajectories of the visually perceived and linguistically described objects are congruent in contrast to if the trajectories are incongruent. The interpretation offered is that in the visual-linguistic match condition, the neural systems used to analyze motion are engaged by the visual stimulus,

and are thus less available for use in sentence comprehension. This leads to interference between the processing of the visual and the language stimuli. While this study may indeed show the behavioral consequences of interference on the neural level, Kaschak et al. (2005) cannot provide conclusive neuro-functional evidence on the basis of their reaction time study.

The visual perception of moving stimuli is known to rely on area MT/V5 in posterior middle temporal cortex (Tootell et al., 1995a; Smith et al., 1998). Importantly, this area has been shown to be responsive both to the actual perception of moving stimuli, as well as to imagery of motion (Goebel et al., 1998). Thus, according to the above mentioned theory it is MT/V5 which is expected to respond both to visually perceived objects in motion and to sentences describing objects in motion.

Those imaging studies investigating visuo-motor language representations to date have focused on the processing of words in isolation (Kable et al., 2002, 2005; Noppeney et al., 2005; Pirog Revill et al., 2008). Typically motion verbs are compared with non-motion verbs (Noppeney et al., 2005; Pirog Revill et al., 2008) or with words describing various types of objects or states (Kable et al., 2002, 2005). For example, Kable and colleagues (2002, 2005) presented participants with a target word and two alternative words, one of which was relatively more semantically related to the target than

the other (e.g., target word: *skipping*; alternatives: *rolling*, *bouncing*). When making decisions about motion verbs, participants showed greater levels of activation in brain areas very proximal to MT/V5 compared to when making the same type of decision about static object words. Typically (as in the case of the studies by Kable and colleagues), greater activation for motion verbs is observed in posterior lateral temporal cortex (PLTC), somewhat anterior to and distinct from MT/V5. Despite the lack of a clean overlap with MT/V5 these results are generally taken to support the notion that retrieving conceptual information about motion (i.e., action) verbs activates visual-motion representations.

One shortcoming in these previous studies is the fact that most have focused on the processing of words in isolation. Specifically, words in isolation may be indeed be associated with motions (or have motion components in their semantic composition), but they are not directly comparable with the visual stimuli which have been used to elicit activation in MT/V5 in the visual domain (see also Wallentin et al., 2005). For example, a typical stimulus used to activate MT/V5 is a display of expanding dots (Tootell et al., 1995a), however neither the word “expand” nor the word “dot” describes the stimulus as well as the combination of “expanding dot display” or better yet the sentence “The dots moved away from the center of the screen.” Thus previous studies using single word stimuli may simply fail to specify enough information about an event to make visual processing possible. In the current two experiments, we therefore presented participants with *sentences* describing objects in motion or objects at rest, and compared the neural correlates of comprehending each sentence type. This design is comparable to what was used in the behavioral study of Kaschak et al. (2005), however it provides direct evidence as to the substrate underlying the previously observed behavioral effect.

A second reason to use sentence stimuli is to create a better match between language stimuli with and without motion content. Specifically, previous studies have frequently contrasted action verbs (i.e., language with motion content) with concrete nouns (i.e., language without motion content). This has led some researchers to question whether observed differences in PLTC might actually reflect differences in grammatical category rather than semantic information (Bedny et al., 2008). On the other hand, while the focus of the aforementioned studies has indeed been on action verbs, Kable et al. (2005) did include a contrast between nouns with implicit motion content (e.g., words denoting manipulable objects) and nouns without motion content (e.g., animals). The results showed that manipulable object nouns elicited greater levels of activation within PLTC (i.e., middle temporal gyrus) than animal nouns. This indicates that differences in PLTC can be found within word categories. Nevertheless, we avoid grammatically dependent activation differences by presenting sentences with similar syntactic structure but different propositional content.

## EXPERIMENT 1: SENTENCES DESCRIBING MOVING VS. STATIC SCENES

### MATERIALS AND METHODS

#### Participants

Twelve participants (six male) aged 22–30 years (mean = 26) participated in this experiment after giving informed written consent.

### Materials

Seventy-two sentences were constructed for the experiment and recorded by a native German speaker (examples can be seen in **Table 1**). Twenty-four of the 72 sentences described items moving toward the participant (SENT<sub>T</sub>); 24 sentences described items moving away from the participant (SENT<sub>A</sub>) and 24 sentences described items which were not in motion (SENT<sub>S</sub>).

In addition three visual stimuli were created. The first of these was a black and white spiral rotated in a clockwise direction to give the illusion of motion toward the participant (VIS<sub>T</sub>); secondly a black and white spiral rotated in a counter-clockwise direction to give the illusion of motion away from the participant (VIS<sub>A</sub>); and thirdly a static image, which was created by scrambling the visual input of the spiral stimuli, which did not create any illusion of motion (VIS<sub>S</sub>). All images were 580 × 580 pixels.

Each of the 72 constructed sentences was presented in conjunction with all three visual stimulus options (VIS<sub>T</sub>, VIS<sub>A</sub>, VIS<sub>S</sub>) yielding 216 critical trials and nine potential trial conditions.

In addition to the critical experimental stimuli, participants received catch trials designed to test their engagement (see below), as well as null events in which no stimulus was presented for the duration of a normal trial.

### Stimulus presentation

Participants lay supine in the scanner. Visual stimuli were presented via 3D glasses onto a virtual computer monitor, and spoken sentence stimuli were presented via headphones. Participants could control a small response box with their right hand (i.e., by pressing one of two buttons). Responses to catch trials were recorded via the response box (see below for more details).

A single critical trial constituted presentation of a sentence in conjunction with visual stimulation. Trial length was 4 s, and the duration of each sentence was approximately 2 s. To enhance the temporal resolution of the acquired signal, trial onset was jittered by 250 ms with respect to the first scan in each trial. In other words, while the first trial was initiated co-incidentally with the beginning of a scan, the second trial and third trials were initiated 250 and 500 ms. after initiation of a scan, respectively. This was done by setting the intertrial interval (ITI) to 4250 ms. Each trial began with presentation of the visual stimulus (VIS<sub>T</sub>, VIS<sub>A</sub>, VIS<sub>S</sub>), which remained visible for the duration of the trial (4 s). An auditory sentence (SENT<sub>T</sub>, SENT<sub>A</sub>, SENT<sub>S</sub>) was presented to participants over headphones 500 ms post onset of visual stimulation.

Following 8% of the trials ( $N = 18$ ) a catch trial was introduced, which did not enter into the functional data analysis. In catch trials participants were asked to indicate whether motion in the preceding

**Table 1 | Examples of the sentences used in Experiment 1 with English translations.**

Condition	Example sentence	English translation
Toward	Das Auto fährt auf Dich zu.	The car drives toward you.
Away	Das Auto fährt von Dir weg.	The car drives away from you.
Static	Das Auto sieht groß aus.	The car looks big.

visual stimulation had been in a clockwise or a counter-clockwise direction. In this manner we ensured that participants processed the visual stimuli.

Experimental stimuli were presented in three blocks of approximately 10 min each with a 1-min pause between blocks (i.e., the purpose of the blocks was only to give participants a rest). Each of the 72 constructed sentences was presented in conjunction with all three visual stimuli options ( $VIS_p$ ,  $VIS_A$ ,  $VIS_s$ ) yielding 216 critical trials. Although sentences were presented three times in the course of the entire experiment, each sentence was presented only once within any given block. The order of visual stimulus presentation in conjunction with a single sentence was balanced across the experiment. The 216 critical trials plus 24 null events (low level baseline condition consisting of a blank screen presented for 4 seconds) were presented in a balanced, pseudorandomized order, such that a single condition was not repeated on more than three consecutive trials, and the probability of each condition following any other condition was matched.

### **FMRI data acquisition**

Scanning was performed using a 3-T MedSpec 30/100 scanner (Bruker, Ettlingen, Germany) and a birdcage head coil. Twenty axial slices (4-mm thickness, 1 mm inter-slice distance, FOV 19.2 cm, data matrix of  $64 \times 64$  voxels, in-plane resolution of  $3 \times 3$  mm) were acquired every 2 s during functional measurements (BOLD sensitive gradient EPI sequence, TR = 2 s, TE = 30 ms, flip angle =  $90^\circ$ , acquisition bandwidth = 100 Hz) with a 3-Tesla Bruker Medspec 30/100 system. Prior to functional imaging T1-weighted MDEFT images (data matrix  $256 \times 256$ , TR = 1.3 s, TE = 10 ms) were obtained with a non-slice selective inversion pulse followed by a single excitation of each slice (Norris, 2000). These images were used to co-register functional scans with previously obtained high-resolution whole head 3D brain scans (128 sagittal slices, 1.5 mm thickness, FOV  $25 \times 25 \times 19.2$  cm, data matrix of  $256 \times 256$  voxels) (Lee et al., 1995).

### **FMRI data analysis**

The functional imaging data was processed using the software package LIPSIA (Lohmann et al., 2001). Functional data were corrected first for motion artifacts using a matching metric based on linear correlation. Data were subsequently corrected for the temporal offset between slices acquired in one scan using a cubic-spline interpolation based on the Nyquist-Shannon Theorem. Low-frequency signal changes and baseline drifts were removed by applying a temporal highpass filter to remove frequencies lower than 1/80 Hz and a spatial Gaussian filter with 10 mm FWHM was applied.

Functional slices were aligned with a 3D stereotactic coordinate reference system using linear registration. The registration parameters were acquired on the basis of the MDEFT and EPI-T1 slices to achieve an optimal match between these slices and the individual 3D reference data set which was standardized to the Talairach stereotactic space (Talairach and Tournoux, 1988). The registration parameters were further used to transform the functional slices using trilinear interpolation, so that the resulting functional slices were aligned with the stereotactic coordinate system. This linear normalization process was improved by a subsequent processing step that performed an additional non-linear normalization

(Thiron, 1998). The transformation parameters obtained from both normalization steps were subsequently applied to the functional data. Voxel size was interpolated during co-registration from  $3 \times 3 \times 4$  mm to  $3 \times 3 \times 3$  mm.

The statistical evaluation was based on a least-squares estimation using the general linear model for serially autocorrelated observations (Worsley and Friston, 1995). The design matrix was generated with a synthetic hemodynamic response function and its first derivative (Josephs et al., 1997; Friston et al., 1998). The model equation, made up of the observed data, the design matrix, and the error term, was convolved with a Gaussian kernel of dispersion of 4 s FWHM.

For each participant three critical contrasts were computed. The first of these served to localize the functional region of interest (fROI) in area MT (see section Functional Region of Interest), known to be important in the processing of visual motion (Tootell et al., 1995b). The second two contrasts (see section Whole Brain Analysis) served to locate areas which were involved in the processing of motion content of sentence materials outside of the predefined fROI. Because individual functional datasets had been aligned to the standard stereotactic reference space, a group analysis based on contrast images could be performed. Single-participant contrast images were entered into a second-level random effects analysis for each of the contrasts. The group analysis consisted of a one-sample *t*-test across the contrast images of all subjects that indicated whether observed differences between conditions were significantly distinct from 0. Subsequently, *t*-values were transformed into *Z*-scores.

### **Functional region of interest**

In order to locate the area MT, a direct contrast between brain activation elicited by moving visual stimuli was compared to activation elicited by static images ( $VIS_{T,A}$  vs.  $VIS_s$ ) in a group average across all participants. In all cases, co-occurring sentence stimuli described static scenes ( $SENT_s$ ). The maximally activated voxel in the posterior middle temporal gyrus plus the 26 voxels adjacent to the peak voxel (ca. 700 mm<sup>3</sup>) in this contrast were identified as the fROI. For each individual participant, time course data was extracted from the voxels in the predefined MT fROI for each Sentence condition ( $SENT_p$ ,  $SENT_A$ ,  $SENT_s$ ) in conjunction with static images ( $VIS_s$ ). In this way, trials belonging to the functional localizer and trials included in the contrast of sentence materials were independent from one another (see also Saxe et al., 2006).

For the analysis of the experimental sentence stimuli, the time course of MR signal intensity was extracted for each individual participant from all voxels within the predefined MT region of interest. The average percent signal change was calculated for each subject and stimulus type, using the average signal intensity during null events as a baseline. Because the fMRI response typically peaks 6 s after stimulus onset, mean percent signal change was calculated for each participant between 4 and 8 s post-stimulus onset.

Mean percent signal change for each participant in each sentence condition was entered into a repeated measure ANOVA with the within subject factor Sentence Meaning (Toward, Away, Static). Within each fROI the two critical comparisons concerned (1) sentences describing motion toward the listener vs. sentences

describing stationary scenes, and (2) sentences describing motion away from the listener vs. sentences describing stationary scenes. Bonferroni correction for multiple comparisons was used in determining the statistical significance of differences between sentence conditions within fROIs.

### Whole brain analysis

In order to identify brain areas outside of MT selectively activated by the motion content of sentences, we compared brain activation elicited by sentences describing motion toward the participant vs. sentences describing static content (SENT<sub>T</sub> vs. SENT<sub>S</sub>) in conjunction with static images (VIS<sub>S</sub>) as well as sentences describing motion away from the participant vs. sentences describing static content (SENT<sub>A</sub> vs. SENT<sub>S</sub>) in conjunction with static images (VIS<sub>S</sub>).

To protect against false positive activation a double threshold was applied by which only regions with a *Z*-score exceeding 3.09 ( $p < 0.001$ , uncorrected) and a volume exceeding 18 voxels (650 mm<sup>3</sup>) were considered (Forman et al., 1995). This non-arbitrary activation size was determined using a Monte Carlo simulation and is equivalent to a corrected significance level of  $p < 0.05$ .

## RESULTS

### BEHAVIORAL RESULTS

Participants responded successfully to the 18 catch trials (mean = 16.3, SD = 2.7). Eliminating two participants who made more than 60% errors improved average performance (mean = 17.4, SD = 1.1), but did not substantially change the functional results. Therefore functional data from all 12 participants were included.

### FUNCTIONAL REGION OF INTEREST (fROI)

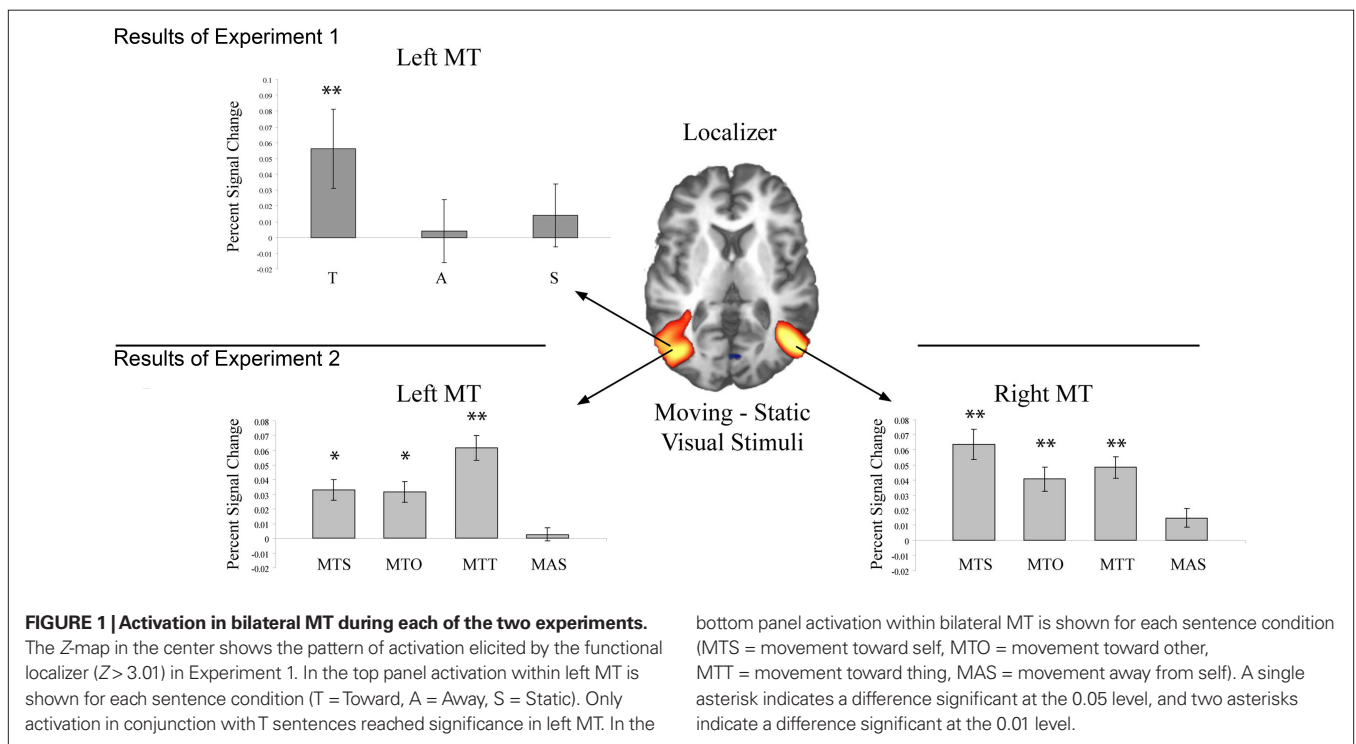
In the left hemisphere MT was identified at Talairach coordinate  $-47, -75, 9$ . In the right hemisphere the peak activated voxel was located at  $43, -67, 9$  (see **Figure 1**).

Mean percent signal change for each participant within left and right MT for the critical sentence conditions was entered into a repeated measure ANOVA with the within subjects factor Sentence Meaning (Toward, Away, Static). In the left hemisphere a main effect of Sentence Meaning was observed [ $F(2,22) = 4.19; p < 0.05$ ]. This reflected significantly more activation for Toward than Static sentences [ $F(1,11) = 9.07, p = 0.01$ ], however no reliable difference in activation level was seen for Away vs. Static sentences [ $F(1,11) < 1$ ]. In the right hemisphere no reliable main effect of Sentence Meaning was observed [ $F(2,22) < 1$ ].

### WHOLE BRAIN ANALYSIS

In addition to the fROI analysis described above, we identified additional areas of the brain which responded selectively to sentence meaning in a whole brain analysis. Relevant brain areas were identified by directly comparing activation elicited by (1) sentences describing motion toward the participant vs. sentences describing static scenes, and (2) sentences describing motion away from the participant vs. sentences describing static scenes.

Sentences describing motion toward the participant vs. sentences describing static images elicited activation within the right posterior superior temporal sulcus, directly superior to the region identified as MT in the right hemisphere by the fROI analysis. In the left hemisphere activation in MT did not pass the cluster-size threshold (see **Table 2**) at the whole brain level, however a smaller area (108 mm<sup>3</sup>) was activated at the level  $p < 0.001$ , uncorrected. Additionally activation was seen in several areas





**Table 2 | Areas showing significantly different activation (A) within posterior middle temporal gyrus (MT) for the functional localizer and (B) whole brain analysis contrasting different sentence types.**

Contrast	Region	Extent (mm <sup>3</sup> )	Z <sub>max</sub>	x	y	z
<b>(A) FUNCTIONAL LOCALIZER</b>						
visual motion – static image	Right posterior middle temporal gyrus (RMT)	7965	5.47	43	–66	9
	Left posterior middle temporal gyrus (LMT)	10854	5.11	–44	–73	9
<b>(B) WHOLE BRAIN ANALYSIS</b>						
static image with toward sentences – static sentences	Right posterior superior temporal gyrus (pSTG)	783	3.77	49	–61	21
	Orbitofrontal median wall	648	3.38	1	65	3
	Middle cingulate gyrus	2079	3.93	10	–7	36
	Posterior cingulate gyrus	4968	3.41	7	–49	30
	Posterior median wall (Cuneus)	1080	3.87	10	–85	42
Static image with away sentences – static sentences	No significant differences					

An activation was regarded significant at a voxel threshold of  $p < 0.001$  and a cluster size  $>640 \text{ mm}^3$  (equivalent to corrected  $p < 0.05$ ). Region, extent (mm<sup>3</sup>), Z-values and Talairach coordinates are reported.

along the median wall including medial prefrontal cortex, middle cingulate gyrus, posterior cingulate gyrus, and cuneus (see **Table 2, Figure 2**).

Sentences describing motion away from the participant vs. sentences describing static images yielded no significantly activated brain regions in the whole brain analysis.

## DISCUSSION

In Experiment 1, we investigated whether comprehending sentences describing an event involving motion reliably activated MT/V5 in contrast to sentences describing a static scene. Previous fMRI research on motion-related language has focused on the processing of action verbs and single nouns (Kable et al., 2002, 2005; Noppeney et al., 2005; Pirog Reville et al., 2008). These studies tend to show activation anterior and dorsal to MT/V5, but not directly within MT/V5. This anterior shift might be due to the fact that an isolated verb or noun gives little specific information about motion parameters. Therefore, participants in the current study listened to sentences describing short events about objects in motion and static scenes.

The most important results from Experiment 1 are as follows. First, the coordinates obtained with the functional localizer task are in accordance with those described in previous literature (Tootell et al., 1995a; Kable et al., 2005). Second, within left MT/V5 activation was modulated by sentence meaning, such that sentences describing objects in motion toward the participant (e.g., “The car drives toward you.”) elicited reliably more activation than sentences describing static scenes (e.g., “The car looks big.”). However, a similar difference was not observed for sentences describing objects in motion away from the participant (e.g., “The car drives away from you.”). Additionally, multiple voxels within the right MT/V5 appear to show modulation by sentence meaning (i.e., Toward > Static, see **Figure 1**), this difference, however, did not reach significance when all voxels in the ROI were considered.

Thus, we see reliable difference in left MT/V5, but only for those motion sentences presented which describe a “toward” motion rather than an “away” motion.

It is possible that the design used in Experiment 1 was suboptimal for detecting subtle changes within MT (for away motion compared to the static scene sentences). Participants were always presented simultaneously with both a language and a visual stimulus. It is conceivable that subtle language-driven changes in MT were weakened by changes elicited in MT throughout the experiment in response to actual moving stimuli. Additionally, posterior temporal cortex (in close proximity to MT) is known to be a multimodal area that receives input from both auditory and visual channels, and this area may be important in the integration of auditory and visual input (Beauchamp, 2005). Thus, the hypothesized general modulation of MT as a function of movement sentences may have been difficult to detect because participants performed some degree of audio–visual integration in all trials. We therefore conducted a second experiment in which presentations of visual and auditory information were kept separate (see further below).

In addition to the ROI analysis, a whole brain analysis was conducted, which revealed significantly more activation in several brain areas for Toward sentences in contrast to Static sentences, but no reliable differences for Away sentences in contrast to Static sentences. In particular Toward sentences showed higher activation in the orbitomedial prefrontal cortex (OMPFC), posterior cingulate cortex (PC) and cuneus (see **Figure 2**).

The cortical midline structures have been implicated in both processing of self-referential stimuli (review see Northoff and Bermpohl, 2004) as well as modulation of visual attention based on cue-induced anticipation (Small et al., 2003). With respect to the first of these points, the OMPFC together with the PC has been seen to play a role in tasks requiring self-reflection, for example indicating whether or not a given word describes oneself vs. indicating whether or not the same word describes another person (Johnson

et al., 2002; Kelley et al., 2002; Kjaer et al., 2002). In the current data, OMPFC and medial posterior parietal cortex are seen to be more activated for sentences describing objects in motion toward the participant than for sentences with no self-referential content (i.e., Static sentences). Activation in these areas may thus reflect the self-referential content of these sentences for participants. The OMPFC is also one of the primary regions to send output to visceromotor structures in the hypothalamus and brainstem, and is thus well-suited to initiate changes in the body (Ongur and Price, 2000). Therefore activation in this area in conjunction with sentences describing objects in motion toward the participant may reflect awareness of the self entering a situation in which action should be initiated (see further below for related discussion).

With respect to visual attention, the cingulate gyrus along with MPFC has been shown to support monitoring and modulation of visual attention (Small et al., 2003; Taylor et al., 2009). Specifically a distinction has been made between anterior and posterior cingulate regions, with anterior cingulate gyrus (AC) active during visual search, spatial working memory and conflict monitoring and PC together with MPFC responsible for the mediation of visual attention based on cue-induced anticipation. It has been suggested that PC is particularly involved when shifts in visual attention are required to efficiently modify behavior (Small et al., 2003). In contrast to the classic distinction between AC and PC, Vogt (2005) proposes a four-region model of cingulate cortex. Specifically, based on the results of monkey and human neurophysiological and cytoarchitectonic studies, cingulate cortex is divided into anterior, middle and posterior cingulate as well as retrosplenial cortex. In this model posterior middle cingulate cortex (pmCC) is involved in skeletomotor orientation while PCC is involved in visuospatial orientation that is mediated through connections to the parietal lobe as well as assessment of self-relevant sensations. In a similar vein, Taylor et al. (2009) have shown that connections between insular cortex and MCC support response selection and skeletomotor body control during monitoring of the environment. The results presented here show greater engagement of PC according to traditional views of cingulate cortex, and greater engagement of pmCC and dorsal PPC according to Vogt (2005) for the comprehension of sentences describing objects in motion toward the listener. We suggest that motion of an item toward oneself in many cases requires monitoring of the environment and subsequent initiation of a motor reaction (i.e., to locate the car moving toward one, or ultimately to move out of the way of the approaching car), whereas objects moving away from the participant or static objects may not require an immediate motor response or modification of behavior. Therefore it is possible that activation in the cortical midline structures for Toward sentences reflect (1) self-relevance of sentences describing approaching objects and (2) preparation of neural systems to identify and behaviorally respond to approaching objects. This interpretation is in line with the role of cingulate cortex proposed by Vogt (2005) and Taylor et al. (2009).

In Experiment 2, in addition to separating visual and language stimuli, we test the hypothesis that self-referentiality is critical in eliciting activity in the cortical midline structures. To this end, participants were presented with sentences describing objects in motion toward not only themselves but also toward inanimate objects (e.g., “The car drives toward the bridge”). If our interpre-

tation is correct, sentences describing objects in motion toward inanimate objects should not activate the cortical midline structures in the same manner as seen for sentences describing objects in motion toward oneself. Moreover, we included a further experimental condition in which sentences describing objects in motion toward other people were presented (e.g., “The car drives toward Maria”). If the observed activation of the cortical midline structures are a reflection of self-referentiality we predict these structures to be active only in the Toward Self sentence, but not in the Toward Other sentences. However, there is recent evidence that the self vs. other distinction in action understanding might be less strict than previously thought (Rizolatti et al., 2001; Wilson and Knoblich, 2005). Under this assumption one might expect an overlap of brain activation for the Toward self and Toward other condition.

## EXPERIMENT 2: REFERENCE TO SELF AND OTHERS IN THE PROCESSING OF MOTION SENTENCES

### MATERIALS AND METHODS

#### Participants

Sixteen participants (eight male) aged 22–27 years (mean = 24.8) participated in this experiment after giving informed written consent.

#### Materials

Sentence material consisted of a total of 192 short sentences. All sentences were constructed in the active voice with a subject, a verb, and prepositional phrase (see examples in **Table 3**). Of the 192 sentences, 120 sentences constituted critical items. These belonged to one of five conditions: Movement Toward Self (MTS), Movement Toward Other (MTO), Movement Toward Thing (MTT), Movement Away from Self (MAS), and No Movement (NM). For each condition 24 sentences were constructed. The remaining 72 non-critical items comprised 24 catch trials (see below) and 48 filler sentences, which were of a form comparable to that of the critical items, but always described objects in motion away from various things. These were included in order to keep the number of occurrences of the words “toward” and “away” balanced.

Catch trials were designed in order to ensure that participants were listening carefully to sentence content. Catch trials were semantically anomalous sentences with the same syntactic form as critical items.

**Table 3 | Examples of the sentences used in Experiment 2 with English translations.**

Condition	Example sentence	English translation
MTS	Das Auto fährt auf Dich zu.	The car drives toward you.
MTA	Das Auto fährt von Dir weg.	The car drives away from you.
MTO	Das Auto fährt auf Maria zu.	The car drives toward Maria.
MTT	Das Auto fährt auf die Brücke zu.	The car drives toward the bridge.
NM	Das Auto sieht groß aus.	The car looks big.

The three visual stimuli created for Experiment 1 were used again. These constituted (1) a black and white spiral rotated in a clockwise direction to give the illusion of motion toward the participant ( $VIS_T$ ); (2) a black and white spiral rotated in a counter-clockwise direction to give the illusion of motion away from the participant ( $VIS_A$ ); and (3) a static image, which was created by scrambling the visual input of the spiral stimuli and did not create any illusion of motion ( $VIS_S$ ). All images were  $580 \times 580$  pixels.

### Stimulus presentation

*Participants lay in the scanner.* Auditory stimuli were presented over headphones and visual stimuli were presented on a virtual monitor seen through 3D glasses. After being instructed and given practice trials, participants listened to all sentence stimuli in a single continuous block (approx. 33 min) and performed a sentence congruency task. Specifically, participants were instructed to respond by pressing a button whenever a semantically incongruent sentence was presented. In this manner, all critical trials remained free of motion artifacts. Following the sentence comprehension block, participants were shown visual stimuli for the purpose of localizing area MT/V5.

In the sentence comprehension block, each trial constituted presentation of a single sentence. The interstimulus interval was approximately 6 s. To enhance the temporal resolution of the acquired signal, a temporal jitter of 500, 1000, or 1500 ms. was inserted into the beginning of each trial.

### FMRI data acquisition

Scanning was performed using a 3-T MedSpec 30/100 scanner (Bruker, Ettlingen, Germany) and a birdcage head coil. Twenty axial slices (4-mm thickness, 1 mm inter-slice distance, FOV 19.2 cm, data matrix of  $64 \times 64$  voxels, in-plane resolution of  $3 \times 3$  mm) were acquired every 2 s during functional measurements (BOLD sensitive gradient EPI sequence, TR = 2 s, TE = 30 ms, flip angle =  $90^\circ$ , acquisition bandwidth = 100 Hz) with a 3-Tesla Bruker Medspec 30/100 system. Prior to functional imaging T1-weighted MDEFT images (data matrix  $256 \times 256$ , TR = 1.3 s, TE = 10 ms) were obtained with a non-slice selective inversion pulse followed by a single excitation of each slice (Norris, 2000). These images were used to co-register functional scans with previously obtained high-resolution whole head 3D brain scans (128 sagittal slices, 1.5-mm thickness, FOV  $25 \times 25 \times 19.2$  cm, data matrix of  $256 \times 256$  voxels) (Lee et al., 1995).

### FMRI data analysis

In order to perform a motion correction, functional volumes were realigned and unwrapped using SPM5. The further processing steps were performed using the software package LIPSIA (Lohmann et al., 2001). To perform a slicetime correction, a cubic-spline interpolation was applied.

In order to align the functional data with a 3D stereotactic coordinate reference system (Talairach and Tournoux, 1988), a linear registration was performed. The registration parameters were acquired using an anatomical reference brain. Hereafter, the registration parameters were used to transform the functional data set to the stereotactic coordinate system, by using a trilinear interpolation. The resulting voxel size was  $3 \times 3 \times 3$  mm.

After registration and normalization, a temporal highpass filter with a cutoff frequency of 1/90 Hz was used for baseline correction of the signal and a spatial Gaussian filter with 6 mm FWHM was applied. The statistical evaluation was based on a general linear regression with pre-whitening (Worsley et al., 2002). Specifically, autocorrelation parameters were estimated from the least-squares residuals using the Yule–Walker equations. These parameters were subsequently used to whiten both data and design matrix. Finally, the linear model was re-estimated using least squares on the whitened data to produce estimates of effects and their standard errors.

### Whole brain analysis: localizer block

As noted before, each individual functional dataset was aligned with the standard stereotactic reference space, so that a group analysis based on the contrast-images could be performed. The single-participant contrast-images were entered into a second-level random effects analysis for each of the contrasts. The group analysis consisted of a one-sample *t*-test across the contrast images of all subjects that indicated whether observed differences between conditions were significantly distinct from 0. The results were corrected for multiple comparisons using cluster-size and cluster-value thresholds obtained by Monte–Carlo simulations using a significance level of  $p < 0.05$  (clusters in the resulting maps were obtained using a *Z*-value threshold of 2.576).

### Functional region of interest (ROI)

The contrast images were also used for regions of interest (ROI) analysis. We tested whether contrast values in ROIs were significantly different from 0 in each experimental condition. Masks for the left and the right MT/V5 were created using the functional images of the localizer run. Three other ROIs were determined using the results of Experiment 1: the orbitomedial prefrontal cortex (OMPFC), posterior cingulate (PC) and cuneus (CUN).

## RESULTS

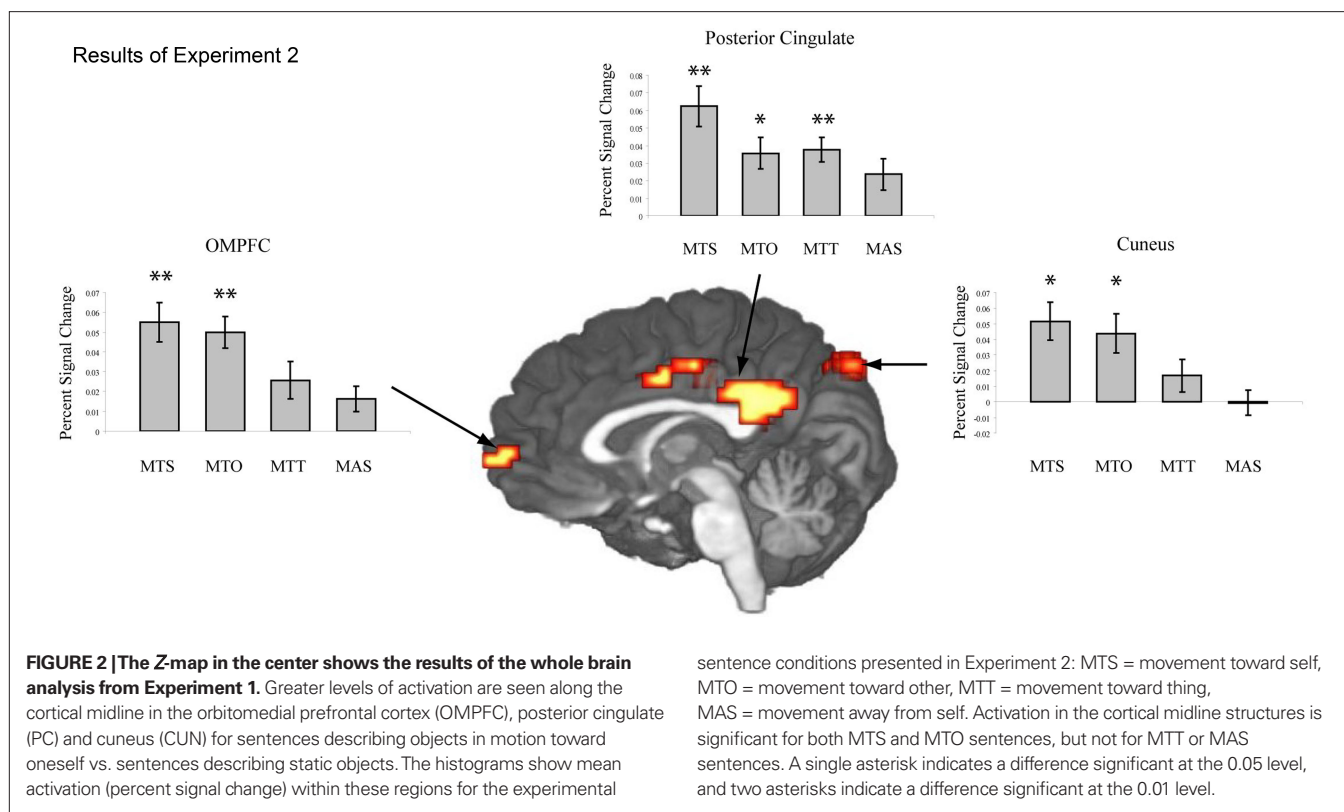
### BEHAVIORAL

Participants showed high performance (93.75%, SD = 7.13) for detection of catch trials (i.e., semantically anomalous sentences), indicating that they were indeed listening to and processing the semantic content of sentences in general.

### FUNCTIONAL REGION OF INTEREST (ROI)

#### Area MT/V5

In the left hemisphere the peak activation in MT/V5 was identified at  $-43, -68, 10$ . In the right hemisphere the peak coordinate was at  $43, -64, 8$  (see **Figure 1**). Planned paired samples *t*-tests showed that in bilateral MT/V5 sentences describing objects in motion toward the participant as well as toward others and toward inanimate objects elicited reliably more activation than sentences describing static scenes (Left MT/V5: MTS-NM:  $t(15) = 2.35, p < 0.05$ ; MTO-NM:  $t(15) = 2.18, p < 0.05$ ; MTT-NM:  $t(15) = 3.69, p = 0.001$ ; Right MT/V5: MTS-NM:  $t(15) = 3.23, p < 0.01$ ; MTO-NM:  $t(15) = 2.47, p < 0.05$ ; MTT-NM:  $t(15) = 3.36, p = 0.005$ ). Sentences describing objects in motion away from participants showed no reliably different activation than sen-



tences describing static scenes in either hemisphere (Left MT/V5: MAS-NM:  $t(15) = 0.28, p > 0.1$ . Right MT/V5: MAS-NM:  $t(15) = 1.23, p > 0.1$ ).

### CORTICAL MIDLINE STRUCTURES

In the OMPFC reliably more activation was seen for sentences describing objects moving toward participants and other people than for sentences describing static scenes (MTS-NM:  $t(15) = 2.82, p < 0.05$ ; MTO-NM:  $t(15) = 3.11, p < 0.01$ ). Sentences describing objects in motion toward inanimate objects and objects in motion away from participants did not activate OMPFC differently than sentences describing static scenes (MTT-NM:  $t(15) = 1.37, p > 0.1$ ; MAS-NM:  $t(15) = 1.21, p > 0.1$ ).

In PC reliably greater activation was observed for sentences describing objects moving toward the participant, toward another person and toward inanimate objects than for sentences describing static scenes (MTS-NM:  $t(15) = 2.7, p < 0.01$ ; MTO-NM:  $t(15) = 1.95, p < 0.05$ ; MTT-NM:  $t(15) = 2.69, p < 0.01$ ). Sentences describing objects in motion away from participants did not activate PC reliably differently than sentences describing static scenes (MAS-NM:  $t(15) = 1.28, p > 0.1$ ).

In the cuneus significantly more activation was recorded for sentences describing objects moving toward the participant and toward another person than for sentences describing static scenes (MTS-NM:  $t(15) = 2.15, p < 0.05$ ; MTO-NM:  $t(15) = 1.75, p = 0.050$ ). There was no reliably different activation for sentences describing objects in motion toward inanimate objects or objects in motion away from participants when compared with sentences describing static scenes (MTT-NM:  $t(15) = 0.79, p > 0.1$ ; MAS-NM:  $t(15) = -0.02, p > 0.1$ ).

### DISCUSSION

The goal of Experiment 2 was twofold. First, we asked whether more conclusive results regarding the role of MT/V5 in processing sentences describing objects in motion could be obtained with more participants and an experimental design separating visual and language stimuli. Secondly, we tested the hypothesis that sentences describing events necessitating reassignment of visual attention and potential modification of behavior to avoid danger elicit more activation than non-threatening sentences in the cortical midline structures (i.e., as seen in Experiment 1).

Using a functional localizer, MT/V5 was identified bilaterally in the lateral posterior middle temporal gyrus. The co-ordinates are very close to those reported in Experiment 1 and also to those reported in previous literature (Tootell et al., 1995a). Within bilateral MT/V5, activation was modulated by sentence content, such that sentences describing objects in motion toward oneself (MTS: “The car drives toward you”) elicited more activation than sentences describing static events (NM “The car looks big”). Activation in MT was not modulated by sentences describing objects in motion away from oneself (MAS: “The car drives away from you”). This suggests that MT/V5 becomes active in comprehending sentences describing approaching but not receding objects. The implications of this are discussed further below.

Cortical midline structures (OMPFC, PC, and cuneus) were reliably more activated for sentences describing objects in motion toward participants than for sentences describing objects in motion away from participants. As discussed previously, the cortical midline structures have been implicated in (1) the processing of self-referential information (particularly OMPFC) and (2) the guiding of visual



attention to support a change in behavior (in particular PC). Thus the results suggest that comprehending self-referential sentences describing an event that would require reallocation of the participant's visual attention (i.e., to detect the approaching object) and a potential change in the participant's subsequent behavior (i.e., to move out of the way of the approaching object) activate those cortical networks known to support execution of these tasks in a natural setting.

In support of this interpretation, sentences describing objects in motion toward inanimate objects (MTT: "The car approaches the bridge.") did not reliably activate the OMPFC. MTT sentences elicited reliable modulation of activation in bilateral MT/V5 and in PC only. These findings are consistent with a network responsible for guiding visual attention toward a moving object (Small et al., 2003), but this network has little overlap with those areas thought to be important in the processing of self-referential information. It is interesting to note here that sentences describing objects in motion toward other things in general (i.e., both toward the self and another object) activate a cortical network involved in the guidance of visual attention and perception of moving stimuli, but sentences describing objects in motion *per se* (e.g., sentences describing objects in motion away from participants) do not. We speculate that in all "toward" conditions, participants may understand a potential need to modify behavior based on the visual scene described. In the case that an object is approaching the participant or an inanimate object, the participant should ultimately change his behavior to avoid the imminent collision. In the case that an object is moving away from the participant, there is no need to modify behavior. Indeed some converging evidence can be found in behavioral studies that argue that looming (i.e., approaching) objects are processed with priority over objects with other motion trajectories (Lin et al., 2008).

Comprehension of sentences describing objects in motion toward other people (MTO: "The car drives toward Mary") elicited activation in those areas also seen to be activated for MTS sentences (i.e., MT/V5, OMPFC, PC, and cuneus). We included this sentence condition in order to test whether activation in the cortical midline structures reflects only self-referential sentence content. The finding that we see more activation in these areas for both self-referential and other-referential sentences suggests that sentences describing objects in motion toward other human beings may be of equal importance to the subsequent behavior of the participant than those sentences describing objects moving toward oneself. The importance may be twofold: the sentence "The car approaches Mary" could either be a warning to a participant that he/she should attempt to avoid hitting Mary with the car or such a sentence could be processed in a manner very similar to those describing objects in motion toward oneself due to the participant's mental representation of the other. In both cases the action plan of the participant must be modified in order to prevent harm to him/herself or the other. Therefore we suggest that activation in OMPFC and cuneus observed during the comprehension of sentences describing objects in motion toward the participant and toward another person reflects awareness that sentence content has potential personal consequences.

## GENERAL DISCUSSION

There are several important findings from this work. First, using a functional localizer, MT/V5 was identified bilaterally in the lateral posterior middle temporal gyrus in both experiments.

The co-ordinates obtained in both studies are very close to those reported in previous literature (Tootell et al., 1995a). Within bilateral MT/V5, activation was modulated by sentence content, such that sentences describing objects in motion toward oneself elicited more activation than sentences describing static events. Whereas activation difference in right MT/V5 only approached statistical significance in Experiment 1, the results of Experiment 2 showed a statistically reliable difference bilaterally. Activation in MT was not modulated by sentences describing objects in motion away from oneself. Taken together, we provide good evidence that sentences describing objects in motion toward participants activate bilateral MT/V5, while sentences describing objects in motion away from participants do not.

Our data are broadly consistent with those reported by Saygin et al. (2010) who also found greater activation in the MT area for sentences conveying motion compared to non-motion sentences. However, Saygin et al. did not report any difference between sentences describing self-motion and those referring to motion of objects and animals. A notable difference between the methodologies used here and in Saygin et al. is that their stimuli were audio-visual, that is, a video tape of a speaker uttering the sentences. Presenting the stimuli in a format closer to a conversational setting may have encouraged greater simulation (in preparation to respond to the speaker). This speculation is consistent with our claim that degree of simulation is subject to top-down modulation.

Another key finding from this work is that brain areas along the median wall, known to be involved in the processing of self-referential stimuli and the direction of attention, are modulated by the self-referential content of experimental sentences. In both experiments sentences describing objects in motion toward the listener elicited greater activation in OMPFC, PC and cuneus than sentences describing static scenes. This pattern was not seen for sentences describing objects in motion away from the listener (Experiments 1 and 2) or for sentences describing objects in motion toward inanimate things (Experiment 2). To summarize, in all sentence conditions in which the propositional content of the sentence could be construed as relevant to the self, activation was elicited in OMPFC and posterior medial parietal cortex. We suggest that these sentences are perceived as more relevant for the participant than sentences describing static scenes or scenes in which the participant's potential actions play no role.

## IMPLICATIONS FOR AN EMBODIED APPROACH TO LANGUAGE COMPREHENSION

The results from the two experiments are partially consistent with the predictions derived from an embodied approach to language comprehension. That is, if language comprehension requires the simulation of sentence content using neural systems also used for perception, action, and emotion, then we should expect activation of MT/V5 during processing of language about visual motion. Indeed we found that activation for sentences describing movement toward oneself, another person, or an object all activated area MT/V5 significantly. However, sentences describing objects in motion away from oneself had no effect on activation level in MT/V5. Thus, although an selective involvement of visual brain areas is seen for the processing of sentences with visual motion content, it cannot be verified that sentences describing objects in motion *per se* activate

MT/V5 in the same manner that actual visual stimuli would. This indicates that higher level visual areas become involved in sentential processing, but only at a late stage and only after sentence meaning has already been at least partially derived.

Our data therefore suggest that some modification of the strong embodiment position that all sentence content is simulated during language comprehension is in order. This notion is in accordance with other research indicating that the simulation view of language comprehension may not be as straightforward as initially thought (e.g., Masson et al., 2008). Specifically, Masson et al. (2008) showed in a behavioral priming study that the type of action-information simulated during comprehension of language denoting manipulable objects differs depending on the sentence context surrounding the manipulable object word. Specifically, they showed that words presented in isolation prime the execution of hand movements related to the functional use of the denoted object (e.g., calculator primes a finger poking action). The same words do not prime execution of hand movements related to moving objects (e.g., calculator does not prime a manual horizontal grasp). However, if manipulable object words are embedded in sentences in which the physical properties of the object become more relevant than the function of the object (e.g., The lawyer kicked the calculator aside.), then the priming effect for functional actions is compromised, while a priming effect for movement actions becomes apparent. This indicates that sensory-motor representations of lexical items are constructed during sentence comprehension. Motor simulation is suggested to draw on whatever experience best captures the relevant properties

of the denoted object in the situation in which it is denoted (e.g., the physical properties of the calculator become more relevant than functional properties in understanding what it means to kick a calculator). There is thus top-down modulation of what kind of action to simulate in conjunction with a manipulable object word. Research to date leaves open what processes underlie this top-down modulation; however the current data suggest that relevance of the described situation for the self is one possible option.

## CONCLUSION

In two fMRI experiments we tested whether area MT/V5, a visual processing area known to be sensitive to moving visual stimuli, is activated by sentences describing objects in motion. The results of our studies show that sentences describing objects in motion toward other entities (people and things) indeed activate higher-order visual brain areas. The results cannot be reconciled fully with a strict embodied interpretation, because sentences describing receding objects did not elicit the same pattern of results in MT/V5. In addition to MT/V5, sentences describing objects in motion toward oneself or another person elicited increased levels of activation along the cortical midline structure. We suggest that this reflects the potential relevance of the information conveyed by the sentence for oneself. The results extend the literature on embodied language processing by showing that higher-order visual areas are also involved in language processing, but they leave open questions concerning the timing with which these areas become activated during sentence comprehension.

## REFERENCES

- Beauchamp, M. (2005). See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Curr. Opin. Neurobiol.* 15, 145–153.
- Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., and Saxe, R. (2008). Concepts are more than percepts: the case of action verbs. *J. Neurosci.* 28, 11347–11353.
- Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., and Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Cogn. Brain Res.* 24, 355–363.
- Forman, S., Cohen, J., Fitzgerald, M., Eddy, W., Mintun, M., and Noll, D. (1995). Improved assessment of significant activation in functional magnetic resonance imaging: use of a cluster-size threshold. *MRM* 33, 636–647.
- Friston, K., Fletcher, P., Josephs, O., Holmes, A., Rugg, M., and Turner, R. (1998). Event-related fMRI: characterizing differential responses. *NeuroImage* 7, 30–40.
- Glenberg, A. M., and Kaschak, M. P. (2002). Grounding language in action. *Psychon. Bull. Rev.* 9, 558–565.
- Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., and Buccino, G. (2008). Processing abstract language modulates motor system activity. *Q. J. Exp. Psychol.* 61, 905–919.
- Goebel, R., Sefat, D., Muckli, L., Hacker, H., and Singer, W. (1998). The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *Eur. J. Neurosci.* 10, 1563–1573.
- Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307.
- Johnson, C., Baxter, L., Wilder, L., Pipe, J., Heiserman, J., and Prigatano, G. (2002). Neural correlates of self-reflection. *Brain* 125, 1808–1814.
- Josephs, O., Turner, R., and Friston, K. (1997). Event-related fMRI. *Hum. Brain Mapp.* 5, 243–248.
- Kable, J., Kan, I., Wilson, A., Thompson-Schill, S., and Chatterjee, A. (2005). Conceptual representations of action in the lateral temporal cortex. *J. Cogn. Neurosci.* 17, 1855–1870.
- Kable, J., Lease-Spellmeyer, J., and Chatterjee, A. (2002). Neural substrates of action event knowledge. *J. Cogn. Neurosci.* 14, 795–805.
- Kaschak, M. P., Madden, C. J., Theriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A., and Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition* 94, B79–B89.
- Kelley, W., Macrae, C., Wyland, C., Caglar, S., Inati, S., and Heatherton, T. (2002). Finding the self? An event-related fMRI study. *J. Cogn. Neurosci.* 14, 785–794.
- Kjaer, T., Nowak, M., and Lou, H. (2002). Reflective self-awareness and conscious states: PET Evidence for a common midline parietofrontal core. *NeuroImage* 17, 1080–1086.
- Lee, J., Garwood, M., Menon, R., Adriany, G., Andersen, P., Truwit, C., and Ugurbil, K. (1995). High contrast and fast three dimensional magnetic resonance imaging at high fields. *Magn. Reson. Med.* 34, 308–312.
- Lin, J., Franconeri, S., and Enns, J. (2008). Objects on a collision path with the observer demand attention. *Psychol. Sci.* 19, 686–692.
- Lohmann, G., Mueller, K., Bosch, V., Mentzel, H., Hessler, S., Chen, L., Zysset, S., and von Cramon, D. Y. (2001). Lipsia – a new software system for the evaluation of functional magnetic resonance images of the human brain. *Comput. Med. Imaging Graph.* 25, 449–457.
- Masson, M., Bub, D., and Warren, C. (2008). Kicking calculators: Contribution of embodied representations to sentence comprehension. *J. Mem. Lang.* 59, 256–265.
- Noppeney, U., Friston, K., and Price, C. (2003). Effects of visual deprivation on the organization of the semantic system. *Brain* 126, 1620–1627.
- Norris, D. (2000). Reduced power multi-slice MDEFT imaging. *J. Magn. Reson. Imaging.* 11, 445–451.
- Northoff, G., and Bermpohl, F. (2004). Cortical midline structures and the self. *Trends Cogn. Sci.* 8, 102–107.
- Ongur, D., and Price, J. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex* 10, 206–219.
- Pirog Revill, K., Aslin, R., Tanenhaus, M., and Bavelier, D. (2008). Neural correlates of partial lexical activation. *Proc. Natl. Acad. Sci.* 105, 13111–13115.
- Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–670.
- Rueschemeyer, S.-A., Brass, M., and Friederici, A. D. (2007). Comprehending prehend: neural correlates of processing verbs with motor stems. *J. Cogn. Neurosci.* 19, 855–865.
- Saxe, R., Brett, M., and Kanwisher, N. (2006). Divide and conquer: a defense of functional localizers. *NeuroImage* 30, 1088–1096.

- Saygin, A., McCullough, S., Alac, M., and Emmorey, K. (2010). Modulation of BOLD response in motion-sensitive lateral temporal cortex by real and fictive motion sentences. *J. Cogn. Neurosci.* 22, 2480–2490.
- Small, D., Gitelman, D., Gregory, M., Nobre, A., Parrish, T., and Mesulam, M. (2003). The posterior cingulate and medial prefrontal cortex mediate the anticipatory allocation of spatial attention. *Neuroimage* 18, 633–641.
- Smith, A. T., Greenlee, M. W., Singh, K. D., Kraemer, F. M., and Hennig, J. (1998). The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). *J. Neurosci.* 18, 3816–3830.
- Talairach, P., and Tournoux, J. (1988). *A Stereotactic Coplanar Atlas of the Human Brain*. Stuttgart: Thieme.
- Taylor, K., Seminowicz, D., and Davis, K. (2009). Two systems of resting state connectivity between the insula and cingulate cortex. *Hum. Brain Mapp.* 30, 2731–2745.
- Thirion, J.-P. (1998). Image matching as a diffusion process: an analogy with Maxwell's demons. *Med. Image Anal.* 2, 243–260.
- Tootell, R., Reppas, J., Dale, A., Look, R., Sereno, M., Malach, R., Brady, T., and Rosen, B. (1995a). Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature* 375, 139–141.
- Tootell, R., Reppas, J., Kwong, K., Malach, R., Born, R., Brady, T., Rosen, B., and Belliveau, J. (1995b). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J. Neurosci.* 15, 3215–3230.
- Vogt, B. (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nat. Rev. Neurosci.* 6, 533–544.
- Wallentin, M., Østergaard, S., Lund, T., Østergaard, L., and Roepstorff, A. (2005). Concrete spatial language: see what I mean? *Brain Lang.* 92, 221–233.
- Wilson, M., and Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychol. Bull.* 131, 460–473.
- Worsley, K., and Friston, K. (1995). Analysis of fMRI time-series revisited – again. *NeuroImage* 2, 173–181.
- Worsley, K., Liao, C., Aston, J., Petre, V., Duncan, G., Morales, F., and Evans, A. (2002). A general statistical analysis for fMRI data. *Neuroimage* 15, 1–15.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 July 2010; accepted: 09 October 2010; published online: 05 November 2010.  
 Citation: Rueschemeyer S, Glenberg AM, Kaschak MP, Mueller K and Friederici AD (2010) Top-down and bottom-up contributions to understanding sentences describing objects in motion. *Front. Psychology* 1:183. doi: 10.3389/fpsyg.2010.00183  
 This article was submitted to *Frontiers in Cognition, a specialty of Frontiers in Psychology*.  
 Copyright © 2010 Rueschemeyer, Glenberg, Kaschak, Mueller and Friederici. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.