Pragmatics in Action: Indirect Requests Engage Theory of Mind Areas and the Cortical Motor Network

Markus J. van Ackeren[1], Daniel Casasanto[2], Harold Bekkering[3], Peter Hagoort[3,4], and Shirley-Ann Rueschemeyer[1,3]

Abstract

■ Research from the past decade has shown that understanding the meaning of words and utterances (i.e., abstracted symbols) engages the same systems we used to perceive and interact with the physical world in a content-specific manner. For example, understanding the word "grasp" elicits activation in the cortical motor network, that is, part of the neural substrate involved in planned and executing a grasping action. In the embodied literature, cortical motor activation during language comprehension is thought to reflect motor simulation underlying conceptual knowledge [note that outside the embodied framework, other explanations for the link between action and language are offered, e.g., Mahon, B. Z., & Caramazza, A. A critical look at the embodied cognition hypothesis and a new proposal for grouding conceptual content. *Journal of Physiology, 102,* 59–70, 2008; Hagoort, P. On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences, 9,* 416–423, 2005]. Previous research has supported the view that the coupling between language and action is flexible, and reading an action-related word form is not sufficient for cortical motor activation [Van Dam, W. O., van Dijk, M., Bekkering, H., & Rueschemeyer, S.-A. Flexibil-

ity in embodied lexical–semantic representations. *Human Brain Mapping*, doi: 10.1002/hbm.21365, 2011]. The current study goes one step further by addressing the necessity of action-related word forms for motor activation during language comprehension. Subjects listened to indirect requests (IRs) for action during an fMRI session. IRs for action are speech acts in which access to an action concept is required, although it is not explicitly encoded in the language. For example, the utterance "It is hot here!" in a room with a window is likely to be interpreted as a request to open the window. However, the same utterance in a desert will be interpreted as a statement. The results indicate (1) that comprehension of IR sentences activates cortical motor areas reliably more than comprehension of sentences devoid of any implicit motor information. This is true despite the fact that IR sentences contain no lexical reference to action. (2) Comprehension of IR sentences also reliably activates substantial portions of the theory of mind network, known to be involved in making inferences about mental states of others. The implications of these findings for embodied theories of language are discussed. ■

INTRODUCTION

Human language is the most sophisticated communication system in the animal kingdom. Specifically, language allows us to encode complex semantic knowledge in a very concise, symbolic way. However, the relationship between symbolic representations and our knowledge of the physical world is still debated. Embodied theories of language postulate that understanding the meaning of words and utterances engages the same systems we use to perceive and interact with the physical world (Barsalou, 1999, 2008; Fischer & Zwaan, 2008; Lakoff & Johnson, 1999). For example, understanding the word "grasp" elicits activation in the cortical motor network. In the embodied literature, cortical motor activation during language comprehension is thought to reflect motor simulation underlying conceptual

knowledge (Barsalou, 1999, 2008). Alternative perspectives, however, such as after-effects of conceptual understanding (Mahon & Caramazza, 2008) and decoding propositional content (Hagoort, 2005), have also been put forth.

Although the functional relevance of the link between language and action remains a topic of debate, there is little doubt that the connection exists: A plethora of previous studies using a variety of experimental techniques have reliably demonstrated that language referring explicitly to actions (e.g., object nouns, action verbs, and action sentences) modulates the activation in the cortical motor network (Glenberg & Gallese, 2011; Willems, Ludovica, D'Esposito, Ivry, & Casasanto, 2011; Van Dam, Rueschemeyer, & Bekkering, 2010; Rüschemeyer, Brass, & Friederici, 2007; Gallese & Lakoff, 2005; Tettamanti et al., 2005; Hauk, Johnsrude, & Pulvermüller, 2004; Glenberg & Kaschak, 2002; Glenberg, 2000). For example, Hauk and colleagues (2004) demonstrated that regions that respond to movement execution with the hand, foot, or mouth will also respond to verbs denoting actions (e.g., pick, kick, and lick). Specifically, the activation patterns revealed a

[1]University of York, UK, [2]New School for Social Research, New York, NY, [3]Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands, [4]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

somatotopic organization of action verbs in the primary and premotor cortices. In another study, Willems and colleagues (2011) applied off-line theta-burst TMS to either the left or the right premotor cortex of their subjects. Subsequently, subjects made lexical decisions on verbs denoting manual actions (e.g., to throw, to write) or nonmanual actions (e.g., to earn, to wander). The results indicated that subjects were faster in making lexical decisions on manual action verbs than nonmanual action verbs when TMS was applied to the left premotor cortex. This finding suggests a functional contribution of the cortical motor system to language understanding. Taken together, these results have been interpreted as evidence that semantic knowledge is instantiated on-line in modality-specific areas in the brain.

There is still much debate regarding the importance of the action-related word form in activating the cortical motor network during language comprehension. Specifically, previous research has challenged the assumption that reading an action-related word form automatically and invariantly activates the cortical motor system. For example, in idiomatic expressions like "kick the bucket," the literal meaning of the utterance denotes an action with the foot. However, the speaker meaning denotes that a person has passed away. Raposo, Moss, Stamatakis, and Tyler (2009) compared idiomatic expressions like these with literal sentences denoting actions and found that the former did not activate the cortical motor network in the same way as literal sentences (but see also Boulenger, Hauk, & Pulvermüller, 2011; Desai, Binder, Conant, Mano & Seidenberg, 2011). This is in line with other results showing less cortical motor activation for nonconventional, nonliteral sentences (e.g., handling the truth) and complex words with motor stems (Rüschemeyer et al., 2007; Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006). These results have supported the view that the coupling between language and action is flexible, and reading an action-related word form is not sufficient for cortical motor activation.

The fact that the relation between word forms referring explicitly to actions and activation of the cortical motor network is not static raises the question whether action words are even necessary to elicit cortical motor activation during language comprehension. Previous research has shown that word meaning activates the cortex in a modality-specific way (Rüschemeyer et al., 2007; Tettamanti et al., 2005; Hauk et al., 2004). However, during natural communication, we do not always express meaning in a literal way. For example, during a conversation, a speaker might mention the fact that "it is quite warm in here." In the appropriate context, this utterance will be interpreted as an indirect request (IR) to open the window, although no explicit action-related word form has been used. Interpreting such a statement as an IR relies on an inference rather than a learned association between form and meaning (Holtgraves, 1994). In other words, the listener has to be aware of the intention of the speaker to interpret the utterance correctly.

In social psychology, making inferences about mental states of others is referred to as having a theory of mind (ToM). Recent research has indicated that areas that are activated when we think about mental states of others in general are also sensitive to pragmatic aspects of language understanding and production (Willems et al., 2010; Sassa et al., 2007). In summary, IRs may refer to an action concept in the absence of an explicit action-related word. Therefore, these utterances allow us to investigate the possibility of triggering activation of the cortical motor network in the absence of action words.

Previous research has shown that action-related word forms are not sufficient for cortical motor activations. The goal of this study goes a step further by investigating whether explicit word forms are a necessary condition for cortical motor activations during language comprehension. Hemodynamic changes in the brain were measured while subjects listened to IRs for actions (e.g., "It is very hot here"). It was hypothesized that (a) the implied meaning of an action should elicit a cortical motor activation. Specifically, we expected activation in areas that are also involved in movement planning, object manipulation and action goals, such as BA 6 and the inferior parietal lobule (IPL; Fogassi et al., 2005). Importantly, these areas have also been associated with understanding actions through language in a large number of studies (Van Dam et al., 2010; Postle, McMahon, Ashton, & de Zubicaray, 2008; Rüschemeyer et al., 2007). Sensitivity of these regions to IRs would suggest that cortical motor activation during language comprehension is not dependent on the presence of action words but could be purely the outcome of a pragmatic inference. In addition, we expected that (b) this inferential step will be reflected by a higher level of activation in classical ToM areas such as the medial pFC (mPFC) and the TPJ (Gallagher & Frith, 2003).

METHODS

Participants

The participants were 16 students from the local university. Three volunteers had to be excluded from the data because of excessive movement, response errors, and health-related problems that were not known to the experimenter before the experiment. The remaining 13 participants were all healthy women between the ages of 18 and 24 years with normal or corrected-to-normal vision and no hearing impairments ($n = 13$; mean age = 21.39 years). All participants were native speakers of Dutch and right-handed. In addition, none of the participants reported any known neurological impairment. Before the experiment, participants gave written informed consent in accordance with the Declaration of Helsinki. For their participation, subjects either received a financial compensation or course credits. The study fell under the legal ethical approval procedure in the Netherlands.

Stimuli

The stimuli consisted of 128 images/visual scenes and 128 spoken sentences. The spoken sentences were recordings from a native speaker of Dutch. The visual scenes were assembled from multiple image search engines on the Web, which are publicly available (e.g., flickr.com; images/google.com).

The stimuli were segregated into 64 unique item sets. One item set comprised four distinct sentence–scene combinations. The four conditions were IR, picture control (PC), utterance control (UC), and picture–utterance control (PUC). Within each item set, only one of these combinations was associated with an IR. Importantly, each request implied a manual action only. The other three sentence–scene combinations could be interpreted as a plausible statement but not as requests. However, the three control conditions were included to control for the effect of the visual scene and the utterance alone. This assumption was tested in an on-line survey before MRI data acquisition. Participants in the questionnaire study ($n = 20$) viewed randomized combinations of one visual scene and one sentence from a single item set. Their task was to decide whether they thought that the interlocutor wanted something from them. The items in the IR condition were interpreted as requests much more often than in the other three conditions (IR: $M = 71\%$; $SE = 4$; PC: $M = 26\%$; $SE = 3$; UC: $M = 18\%$; $SE = 2$; PUC: $M = 14\%$; $SE = 2$). This difference was significant at an alpha level of .05, $F(1.92, 36.33) = 110.31, p < .001$, partial $\eta^2 = .85$.[1]

In addition, a localizer task was used to identify (a) regions that were specifically activated when thinking about another person's believes and desires (ToM) and (b) regions that were involved in simple hand actions. The ToM network was identified with the localizer during time windows at which the subjects read a story and subsequently judged statements about the story. The stories required either an inference on the physical state of an object (false photograph) or another person's belief (false belief). Each participant saw 24 physical state and 24 mental state stories. The stimuli were introduced by the Saxe lab[2] (Dodell-Feder, Koster-Hale, Bedny, & Saxe, 2010; Saxe & Kanwisher, 2003) and translated into Dutch for the present experiment. Regions that were involved in simple hand actions were localized during time windows in which subjects produced button presses with their right index and middle finger. Hand regions were targeted because the indirect requests required simple hand actions as well (Figure 1).

Stimulus Presentation

For both the main task and the functional localizer, subjects lay supine in the scanner. All visual materials were presented via a projector outside the scanner. Participants viewed the screen via a nonmagnetic mirror. The auditory
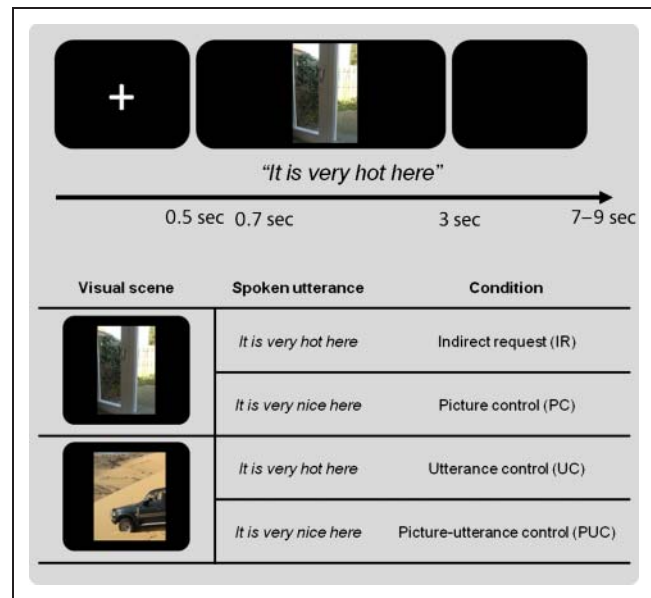

Figure 1. Illustration of the fMRI paradigm for IRs. The top half shows the time course of the experiment. On each trial, a fixation cross was presented for 500 msec, followed by a visual scene. The utterance was presented auditorily, 200 msec after picture onset. Each trial lasted exactly 3000 msec. On 15% of the trials, the subject was prompted to indicate whether the utterance was a request. The bottom half depicts the design of the study with a single item set. The conditions are IR, PC, UC, and PUC.

stimuli were presented via nonmagnetic headphones that also dampened the noise from the scanner. Before the start of the experiment, the volume of the headphones was adjusted to the subjects convenience. Participants' responses to the tasks were recorded via a nonmagnetic button box inside the scanner.

The implicature task was an epoch-related design in which each participant saw two independent sentence–scene combinations from each item set, resulting in a total number of 128 stimuli per participant (32 items per condition). The items were individually pseudo-randomized in such a way that the same condition was never presented more than twice in a row. Additionally, 16 null events were included in the design. To maintain participants' vigilance and ensure that they processed the stimuli more deeply, 15% of the experimental trials were accompanied by a catch question ("Do you think that the person made a request?"), to which participants could respond with a button press indicating either a "yes" or "no" response. To make the trial onset unpredictable for the subject and to enhance the resolution of the time window within a trial, the intertrial interval was randomly jittered in a range of 4000–6000 msec ($M = 5000$ msec). The trial began with the presentation of a fixation cross for 500 msec, followed by the visual scene. After 200 msec, the sentence stimulus was presented (mean duration = 1357 msec). At the sentence offset, a variable interval filled the remaining time so that every picture presentation lasted 2300 msec. Thus, every trial lasted exactly 3000 msec.

Participants were instructed to listen to the sentences carefully and decide whether they think the person wanted something from them or not while listening. Before the actual start of the experiment, there was a practice run outside the scanner.

After the implicature task, subjects proceeded with the localizer task. The procedure of this task is described in Dodell-Feder et al. (2010). Also for this task, there was a practice run outside the scanner, before the experiment.

fMRI Data Acquisition

MRI data acquisition was performed on a Siemens Magnetom Trio scanner (Siemens Medical System, Erlangen, Germany) with a magnetic field strength of 3 T. The functional scans for the implicature task and the localizer were acquired using a multiecho gradient pulse sequence (repetition time = 2390 msec; echo time = 9.4, 21.17, 32.94, 44.71, and 56.48 msec; flip angle = 90°). Each volume consisted of 31 transversal slices with a thickness of 3 mm. The voxel resolution was 3.5 mm × 3.5 mm × 3.5 mm.

After the collection of functional data, a structural scan was performed for each individual participant. The image was a T1-weighted 3-D MPRAGE sequence comprising 192 sagittal slices (repetition time = 2300 msec; echo time = 3.03 msec; slice thickness = 1 mm).

fMRI Data Analysis

The raw MR images were preprocessed and analyzed using the Matlab toolbox SPM8 (Statistical Parametric Mapping, www.fil.ion.ucl.uk/spm). Before the analysis, the first six volumes were excluded to control for T1 equilibration effects. Six movement parameters (three translations and three rotations) were extracted from the first echo of each volume and subsequently used to correct for small head movements in all five echoes of each volume. Subsequently, all five echoes were combined into a single volume using a weighted average. To correct for delays in slice timing during image acquisition, the time courses of each voxel were realigned toward slice 16. After segmentation into gray and white matter, images were normalized to a standard EPI template within Montreal Neurological Institute (MNI) space and resampled at an isotropic voxel size of 3 mm. Lastly, the images were convolved with a Gaussian smoothing kernel with 8-mm FWHM. To correct for slow drifts in the signal, a high pass filter was applied at 128 sec.

The combined and preprocessed time series of the implicature task was analyzed as an event-related design (epoch = 1.5 sec) on a subject-by-subject basis. Within a general linear model framework, each condition was convolved with a canonical hemodynamic response function (HRF) and used as a regressor. In addition, the movement parameters from the realignment algorithm, time, and dispersion derivatives were included as effects of no interest.

A single contrast comparing IR versus three control (PC, UC, and PUC) conditions was generated for each participant individually. Because the images from each subject had been aligned to standard MNI space, a second level random effects analysis could be performed at the group level. The critical contrasts from all participants were included in the model, and a group analysis was performed using a one-sample t test. To control for multiple comparisons, a cluster extend threshold was determined using a Monte Carlo simulation with 1000 iterations (Slotnick, Moo, Segal, & Hart, 2003). The simulation results indicated that a cluster with $p < .0001$ (uncorrected) and a cluster size $k > 12$ (324 mm^3) was significant at $p < .01$ (corrected).

The localizer task was used to identify areas that were selectively active during inferences about mental states and hand movements. To extract the signal, which was related to the ToM network, the images were analyzed as a block design. Each block was defined as the period of time from the onset of the story to the offset of the statement. Subsequently, this time window was convolved with a canonical HRF. Only movement parameters were included as effects of no interest. A contrast comparing the false belief stories versus the false photograph stories was created for each subject individually. Subsequently, a random effects group analysis was conducted on the individual participant contrast images using a one-sample t test.

Additionally, button presses with the right hand were analyzed to identify regions that are involved in action execution. Manual button presses were analyzed as an event-related design. The moment a subject pressed a button was modeled with a canonical HRF and used as a regressor. In addition, the movement parameters from the realignment algorithm, time, and dispersion derivatives were used as effects of no interest.

ROI Analysis

ROI analyses were conducted to investigate whether the condition in which the utterance and the visual context formed an implicature exerted a stronger effect on (a) the neural motor network that is involved in hand actions and (b) the ToM network. Four ROIs for the ToM network were defined as the clusters in the whole-brain analysis that were sensitive to the contrast false belief versus false photograph (Table 1). With respect to the action network, there were very strong anatomical hypotheses. Therefore, the contrast image for hand actions (Action > 0) from the second-level whole-brain analysis was masked with cytoarchitectonically defined probability maps of BA 6 (Geyer, 2004) and left and right IPL/PF (Caspers et al., 2006, 2008; Table 2). The former will be referred to as IPL in the remainder of this discussion. Subsequently, MNI coordinates for peak values within the largest active cluster were used to create spheres of 6-mm radius using the ROI toolbox Marsbar (Brett, Anton, Valabregue, & Poline, 2002).

The ROIs from the localizer task were interrogated with respect to the four conditions (IR, PC, UC, PUC) from the implicature task. Percent signal change was extracted and averaged within each participant. Thus, for each of the 13 participants in our study, there were four values. With these four conditions, ANOVAs with repeated measures were conducted for each ROI.

RESULTS

Behavioral Results

Behavioral responses to catch trials were analyzed to test whether subjects responded as predicted by the questionnaire study. First, a one-sample t test was conducted on the percentage of correct responses to assess whether participants were able to do the implicature task. This assumption was confirmed, $t(12) = 9.64$, $p < .001$, $M = 80\%$, $SE = 3$. Subsequently an ANOVA with repeated measures was conducted on the RTs to test for differences in task difficulty between the conditions (IR, PC, UC, PUC). This test was significant, $F(2.14, 24.43) = 3.56$, $p < .05$, partial $\eta^2 = .23$. Planned comparisons revealed participants responded to the IR condition faster than to the PC [$F(1, 12) = 13.13$, $p < .005$, partial $\eta^2 = .52$; IR: $M = 045$ msec, $SE = 86$ msec; PC: $M = 1299$ msec, $SE = 116$ msec], the UC [$F(1, 12) = 6.1$, $p < .05$, partial $\eta^2 = .34$; UC: $M = 1193$ msec, $SD = 91$ msec], and PUC [$F(1, 12) = 5.05$, $p < .05$, partial $\eta^2 = .3$; PUC: $M = 1314$ msec, $SE = 131$ msec] condition. The fact that participants responded faster, giving positive responses, is not surprising; however, it shows that recognizing IR was not more difficult than recognizing statements. Lastly, an ANOVA with repeated measures tested whether participants recognized requests more often in the IR condition. The results replicated the findings from the questionnaires, $F(2.74, 32.93) = 49.76$, $p < .001$, partial $\eta^2 = .81$. Specifically, requests were more often identified in the IR condition than in the PC [$F(1, 12) = 113.81$, $p < .001$, partial $\eta^2 = .91$; IR: $M = 83\%$, $SE = 4$; PC: $M = 19\%$, $SE = 4$], UC [$F(1, 12) = 71.44$, $p < .001$, partial $\eta^2 = .86$; UC: $M = 34\%$, $SE = 6$], and PUC [$F(1, 12) = 101.57$, $p < .001$, partial $\eta^2 = .9$; PUC: $M = 12\%$, $SE = 7$] condition.

fMRI Results

ToM Localizer

A whole-brain analysis on the ToM localizer was conducted, in which the story and the statement were modeled as one block. The pattern of results replicated previous findings in English (Dodell-Feder et al., 2010). That is, regions that are part of the ToM network showed a stronger BOLD response for false belief stories than for false photograph stories (Table 1). These were clusters in the left and right TPJ, the left precuneus, and the mPFC. These clusters were interrogated in the subsequent ROI analysis of the ToM network.

Action Localizer

To identify regions that were sensitive to action preparation or execution, a whole-brain analysis was conducted on the localizer. Specifically, the moment of a button press, convolved with a canonical HRF, was used as regressor. Because there were very specific anatomical predictions, the image was masked with an anatomical map of BA 6 and bilateral IPL. Table 2 shows the peak activations within these regions. Specifically, activation peaks were found in the left superior frontal gyrus (SFG), left precentral gyrus, and left medial SFG. In functional terms, these peaks are located within the left premotor cortex and the left pre-SMA. In addition, there were peak activations in the left and right IPL, overlapping with the supramarginal gyrus. The peak activations in the frontal motor regions (left SFG, left PCG, and left pre-SMA) as well as the strongest activation peak in left and right IPL were used to create 6-mm spheres for the subsequent ROI analysis.

Whole-brain Analysis

An overview of significant peak activations in the whole-brain analysis of the implicature task is depicted in Table 3. On the medial surface of the brain, a widespread cluster of activation was found in the mPFC and the insular cortex, extending into the OFC. Additionally, regions in the posterior and middle cingulate cortex were sensitive to

Table 1. Brain Regions from Whole-brain Analysis of the Localizer Task Showing Significantly More Activation during False Belief versus False Photograph Stories ($p < .0001$, $k > 12$)

Region	Cluster Level Extent (Voxels)	Peak Voxel Level		MNI Coordinates		
		t	$equivZ$	x	y	z
Precuneus	491	14.56	5.83	−3	−58	22
Left mPFC	23	8.33	4.71	−6	50	40
Right TPJ	115	7.88	4.59	57	−55	22
Left TPJ	156	8.93	4.86	−51	−49	31

Only the largest peak voxel per cluster is depicted.

Table 2. Brain Region's Whole-brain Analysis of the Localizer Task that Were Significantly Active during Button Presses with the Right Hand ($p < .0001$, $k > 12$)

Region	Cluster Level Extent (Voxels)	Peak Voxel Level		MNI Coordinates		
		t	$equivZ$	x	y	z
BA 6	458					
Left SFG		12.66	5.56	−24	−1	67
Left precentral gyrus		11.84	5.43	−24	−10	64
Left medial SFG		10.90	5.27	−6	23	58
Left inferior parietal cortex (PF)	113					
Left inferior parietal lobule		13.21	5.64	−48	−46	37
Right inferior parietal cortex (PF)	40					
Right supramarginal gyrus		7.00	4.34	66	−40	28

The results were restricted to anatomically defined neural motor regions (BA 6, bilateral IPL).

the manipulation. On the lateral surface, there were clusters of activation around the left and right TPJs. However, the latter was much more dominant in the left hemisphere. Lastly, both thalami showed selective activation to IR.

ROI Analysis

In the ROI analysis, ROIs from the ToM and action localizer were interrogated with respect to the average percent signal change in the implicature task. All inferential statistics in this section were evaluated at an alpha level of $p < .05$. ANOVAs with repeated measures were corrected for violations of sphericity using Greenhouse–Geisser correction.

ToM ROIs

ANOVAs with repeated measures were conducted for each ROI from the ToM localizer (mPFC, precuneus, left TPJ, and right TPJ) to test whether these regions were sensitive to IRs. Interrogation of the mPFC revealed a significant main effect of Condition, $F(2.39, 28.72) = 4.68$, $p < .05$, partial $\eta^2 = .28$. Planned comparisons indicated that the average percent signal change was higher in the IR condition than in the three control conditions [IR vs. PC: $F(1, 12) = 5.39$, $p < .05$, partial $\eta^2 = .31$; IR: $M = .11$, $SE = .04$; PC: $M = .01$, $SE = .05$; IR vs. UC: $F(1, 12) = 14.14$, $p < .005$, partial $\eta^2 = .54$; UC: $M = .01$, $SE = .04$; IR vs. PUC: $F(1, 12) = 16.29$, $p < .005$, partial $\eta^2 = .58$; PUC: $M = −.04$, $SE = .04$; Figure 2A]. ANOVAs on the left TPJ also revealed a significant main effect of Condition, $F(2.35, 28.21) = 10.47$, $p < .001$, partial $\eta^2 = .47$. Specifically, planned comparisons showed that average percent signal change was higher in the IR condition than in the control conditions [IR vs. PC: $F(1, 12) = 15.19$, $p < .005$, partial $\eta^2 = .56$; IR: $M = .26$, $SE = .05$; PC: $M = .13$, $SE = .06$; IR vs. UC: $F(1, 12) = 12.91$, $p < .005$, partial $\eta^2 = .52$; UC: $M = .14$, $SE = .06$; IR vs. PUC: $F(1, 12) = 27.02$, $p <$

$.001$, partial $\eta^2 = .69$; PC: $M = .03$, $SE = .06$; Figure 2A]. In addition, there was a significant main effect in the right TPJ, $F(1.92, 22.98) = 5.04$, $p < .05$, partial $\eta^2 = .3$. Planned comparisons revealed the same effects as in the left TPJ, albeit that the contrast with the UC condition was only marginally significant [IR vs. PC: $F(1, 12) = 11.91$, $p < .01$, partial $\eta^2 = .5$; IR: $M = .17$, $SE = .05$; PC: $M = .06$, $SE = .1$; IR vs. PUC: $F(1, 12) = 8.6$, $p < .05$, partial $\eta^2 = .42$; PUC: $M = −.03$, $SE = .05$; IR vs. UC: $F(1, 12) = 3.33$, $p = .09$; UC: $M = .10$, $SE = .06$]. Lastly, the analysis of the precuneus did not reach significance, $F(1.78, 21.33) = 2.06$, $p > .1$.

Action ROIs

ANOVAs with repeated measures were conducted for each of the five cytoarchitectonically and functionally defined ROIs from the action localizer (left SFG, left PCG, left pre-SMA, left IPL, and right IPL) to estimate the sensitivity of these areas to IRs. Interrogation of the left pre-SMA revealed a significant effect of Condition, $F(2.21, 26.5) = 5.92$, $p < .01$, partial $\eta^2 = .33$. Planned comparisons indicated that the level of activation in the IR condition was higher than in the three control conditions [IR vs. PC: $F(1, 12) = 9.86$, $p < .01$, partial $\eta^2 = .45$; IR: $M = .12$, $SE = .05$; PC: $M = .02$, $SE = .05$; IR vs. UC: $F(1, 12) = 5.49$, $p < .05$, partial $\eta^2 = .31$; UC: $M = .01$, $SE = .06$; IR vs. PUC: $F(1, 12) = 16.94$, $p < .005$, partial $\eta^2 = .59$; PUC: $M = −.02$, $SE = .05$; Figure 2B]. The analysis of the left IPL yielded a significant effect of Condition, $F(2.69, 32.29) = 9.63$, $p < .001$, partial $\eta^2 = .45$. Planned comparisons indicated that the average percent signal change was higher in the IR condition than in the control conditions [IR vs. PC: $F(1, 12) = 14.65$, $p < .005$, partial $\eta^2 = .55$; IR: $M = .16$, $SE = .04$; PC: .06, $SE = .05$; IR vs. UC: $F(1, 12) = 15.18$, $p < .005$, partial $\eta^2 = .56$; UC: $M = .06$, $SE = .05$; IR vs. PUC: $F(1, 12) = 30.09$, $p < .001$, partial $\eta^2 = .72$; PUC: $M = .03$, $SE = .04$; Figure 2B]. Also, there was a main effect of Condition in the right IPL, $F(2.68, 32.17) = 5.02$, $p < .01$, partial $\eta^2 = .3$.

Table 3. Brain Regions from Whole-brain Analysis of the Implicature Task Showing Significantly More Activation during IRs versus Controls (PC, UC and PUC; $p < .0001$, $k > 12$)

Region	Cluster Level Extent (Voxels)	Peak Voxel Level t	equivZ	MNI Coordinates x	y	z
Medial frontal cortex	323					
Right ACC		9.08	4.89	15	32	25
Left ACC		8.40	4.73	−3	44	19
Right ACC		8.02	4.63	6	41	19
Right middle frontal gyrus	52	8.76	4.82	21	56	28
Right pars opercularis	21	6.92	4.32	45	17	13
Left precentral gyrus	13	6.76	4.26	−51	2	46
Right insular cortex	69					
Right insula lobe		7.77	4.56	39	26	−5
Pars triangularis		6.70	4.24	42	35	−2
Right putamen		6.02	4.01	24	23	−8
Left insular cortex	139					
Left insula lobe		8.08	4.65	−36	20	−5
Pars orbitalis		7.59	4.51	−39	23	−14
Thalamus	37					
Right thalamus		7.04	4.35	9	−7	7
Left thalamus		6.54	4.19	−6	−19	13
Right middle temporal gyrus	12	6.53	4.19	51	−49	13
Right posterior middle temporal gyrus	15	7.46	4.48	45	−31	−5
Posterior cingulate cortex	232					
Posterior cingulate cortex		9.03	4.88	−6	−37	22
Left middle cingulate cortex		8.17	4.67	−12	−40	37
Cingulate gyrus		8.09	4.65	6	−34	25
Left TPJ	179					
Left supramarginal gyrus		9.07	4.89	−60	−46	28
Left angular gyrus		9.07	4.89	−48	−49	31
Right TPJ	19					
Right supramarginal gyrus		6.46	4.16	63	−43	25
Right superior temporal gyrus		5.81	3.94	63	−52	22
Right supramarginal gyrus		5.49	3.81	51	−40	28
Right precuneus	53					
Right precuneus		11.16	5.31	6	−67	37
Right precuneus		6.42	4.15	18	−52	37
Left precuneus	21	9.90	5.07	−9	−64	37

Figure 2. ROIs were interrogated with respect to the conditions IR, PC, UC, and PUC. The image shows all ROIs, superimposed on a brain template. The bar diagrams illustrate mean percent signal change for each condition. The error bars depict the standard error. (A) Green ROIs show regions from the ToM localizer (mPFC and TPJ). (B) Red ROIs refer to regions that were activated during action execution (pre-SMA and bilateral IPL).
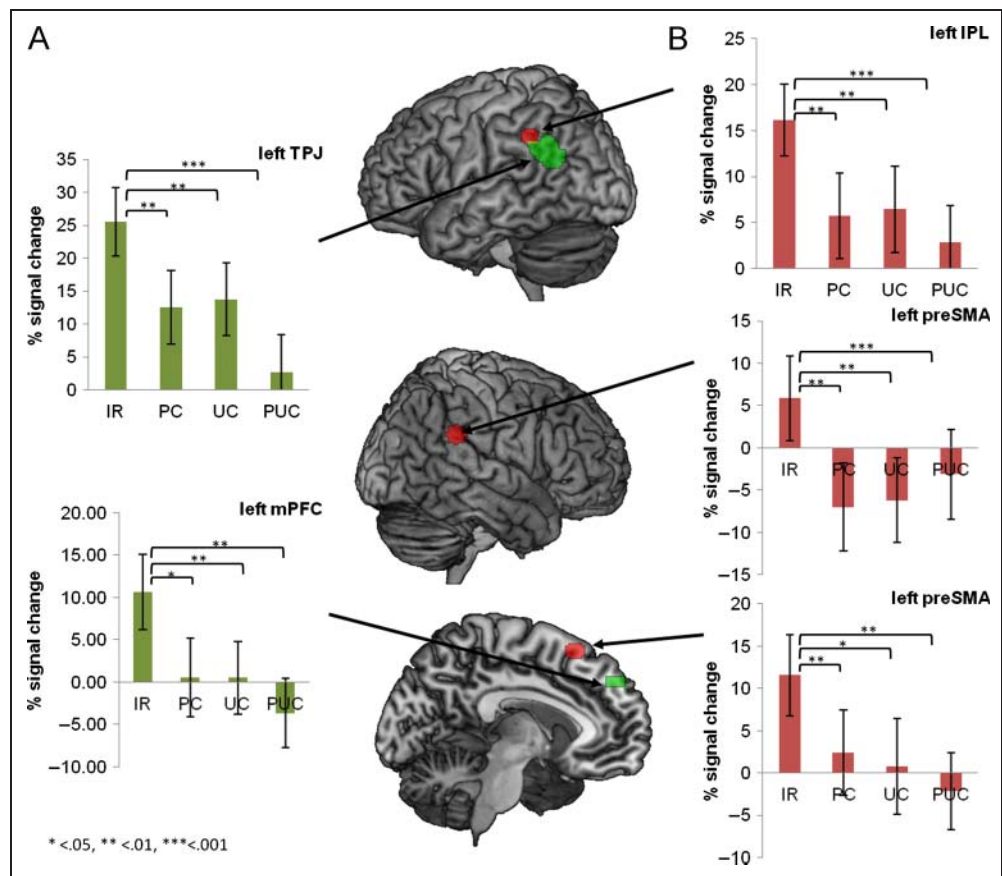
Planned comparisons showed that average percent signal change was higher in the IR condition than in the control conditions [IR vs. PC: $F(1, 12) = 13.77, p < .005$, partial $\eta^2 = .53$; IR: $M = .06, SE = .05$; PC: $M = -.07, SE = .05$; IR vs. UC: $F(1, 12) = 10.58, p < .01$, partial $\eta^2 = .47$; UC: $M = -.06, SE = .05$; IR vs. PUC: $F(1, 12) = 6.2, p < .05$, partial $\eta^2 = .34$, PUC: $M = -.03, SE = .05$; Figure 2B]. However, ANOVAs with repeated measures, investigating the effect of condition in the left SFG and the left PCG did not reach significance, $F(2.9, 34.81) = 1.34, p > .1$ and $F(1.77, 21.2) = .94, p > .1$, respectively.

DISCUSSION

Previous research has demonstrated that language referring explicitly to actions (e.g., action verbs, nouns referring to tools, action sentences) reliably activates cortical motor areas in the brain (Willems et al., 2011; Van Dam et al., 2010; Rüschemeyer et al., 2007; Tettamanti et al., 2005; Hauk et al., 2004). In the current study, we asked whether language that is used to refer implicitly to actions (i.e., without any explicit lexical reference to action) shows a similar pattern of activation. In other words, can cortical motor activation be triggered by expressions without any lexical item that refers to an action? To this end, participants were presented with spoken utterances, some of which could be understood as IRs for actions

(e.g., "it is very hot here," as a request that one open the window) and some of which were simply descriptions of visual scenes (e.g., "it is very hot here," in the context of a desert scene). The results indicate (1) that comprehension of IR sentences activates cortical motor areas reliably more than comprehension of sentences devoid of any implicit motor information. This is true despite the fact that IR utterances contain no lexical reference to action. (2) Comprehension of IR sentences also reliably activates substantial portions of the ToM network, known to be involved in making inferences about mental states of others (Frith & Frith, 2005, 2010; Gallagher & Frith, 2003; Saxe & Kanwisher, 2003). The implications of these findings for embodied theories of language are discussed below.

IRs and the Cortical Motor System

IR sentences activated areas within the larger cortical motor system significantly more than sentences in any of the three control conditions (Figure 2B). This activation pattern was assessed in two ways: (1) in an ROI analysis and (2) in a whole-brain analysis.

In the ROI analysis, voxels in BA 6 and bilateral IPL that were also sensitive to finger movements during the localizer task (button presses) were identified as ROIs. These regions comprised voxels in the left premotor cortex, bilateral IPL, and pre-SMA. Interrogation of these ROIs

with regard to the four language conditions showed that the bilateral IPL and pre-SMA were sensitive to the implicit motor content in IR sentences. In other words, bilateral IPL and pre-SMA showed significantly greater activation for IR sentences than for sentences in any of the three control conditions. In the following paragraphs, we discuss the potential role of the areas targeted by the ROI analysis to processing IR sentences.

The inferior parietal lobe is a sensorimotor area that is often associated with the representation of action goals (Fogassi et al., 2005). Fogassi and colleagues (2005) addressed this hypothesis using single-cell recordings in nonhuman primates. The authors found that a different set of neurons fired when a monkey grasped food to put it in a container than when it was going to eat the food. In addition, some neurons showed the same pattern during action observation. These results provide strong evidence that the neurons in the IPL are sensitive to the goal of an action. Recently, Aziz-Zadeh and Damasio (2008) have argued that the IPL encodes the set of sensorimotor events that coincide with action execution. For example, a movement such as grasping a cup will elicit somatosensory, proprioceptive, and visual feedback. The signals from these three different sources are aligned in time and therefore more likely to be associated in the brain (Aziz-Zadeh & Damasio, 2008). The integration of sensory information and action is particularly important for the functional manipulation of tools. Numerous neuropsychological studies have associated lesions in the IPL with a disability in manipulating objects and tools in a meaningful way (apraxia; for a review, see Wheaton & Hallett, 2007). The IPL is also consistently activated in studies investigating the comprehension of action language (Rueschemeyer, van Rooij, Lindemann, Willems, & Bekkering, 2010). In a study using functional imaging, Rueschemeyer, van Rooij, et al. (2010) found that the IPL was more sensitive to words denoting functionally manipulable objects (e.g., cup) as compared with volumetrically manipulable words (e.g., bookend). Thus, the type of action information that is instantiated in the IPL (i.e., information about complex action plan and how to manipulate objects) appears to be relevant both for executing actions and for processing conceptual information about tools and actions through language.

The pre-SMA is usually associated with executive aspects of motor control (Rushworth, Walton, Kennerley, & Bannerman, 2004; Picard & Strick, 2001). Specifically, Rushworth and colleagues (2004) have suggested that the pre-SMA is involved in selecting and changing between task-relevant action sets. That is, the selection of a specific response from a set of possible responses to a sensory stimulus. This idea is supported by the finding that changing an action set is perturbed if rTMS is applied to the medial SFG (Willems et al., 2011). Although the pre-SMA is not consistently activated in studies investigating the comprehension of action language, this is certainly not the first time that this area has been observed (Rueschemeyer,

van Rooij, et al., 2010; Postle et al., 2008). For example, Postle and colleagues investigated the sensitivity of the cortical motor system (BA 6 and BA 4) to action verbs and found that the pre-SMA in BA 6, but not M1, is sensitive to action verbs. In the current study, we argue that once a participant has understood that a request for action is being made, he or she must evaluate what action is being requested. The pre-SMA could potentially reflect the attempt of the listener to select the best action alternative from the set of possible actions one could perform in any given situation.

In the whole-brain analysis, activation elicited by IR sentences compared with the three control conditions was assessed. The results demonstrate that IR sentences activated a fronto-parietal network, comprising the posterior middle frontal gyrus, left precentral gyrus, and several regions in the bilateral inferior parietal lobe, most notably the supramarginal gyrus. These results are largely consistent with the pattern observed in the ROI analysis and thus provide converging evidence for the involvement of neural motor areas in the processing of IR sentences compared with sentences devoid of any motor content.

Our results indicate that language material devoid of explicit action content activates neural motor areas in the brain if presented in a situation in which reference to an action is communicatively implied. This result marginalizes the role of specific lexical items for cortical motor activation in showing that action-related words are not necessary for activating the cortical motor system during language comprehension. Yet, this does not mean that the motor system is not involved in language processing. Rather, the present findings suggest that motor areas might be involved in language comprehension in a much more complex way than theorists imagined just a decade ago (Pulvermüller, 1999, 2005). However, this idea is in line with more recent approaches, suggesting that the cortical motor system might contribute to language understanding, but other sources of information are also important (Kiefer & Pulvermüller, 2012; Meteyard, Rodriguez Cuardrado, Bahrami, & Vigliocco, 2012; Van Dam, van Dijk, Bekkering, & Rueschemeyer, 2011).

IRs and ToM

To understand the speaker meaning of an utterance, the listener needs to infer the communicative intent of the speaker (Holtgraves, 1994; Grice, Cole, & Morgan, 1975). Recent neuroimaging studies have addressed the relationship between language and communication (Willems et al., 2010; Sassa et al., 2007). Evidence from these studies suggests that the brain regions that are modulated by the communicative demands of an utterance overlap with classical ToM regions, but not language regions. For example, Willems and colleagues (2010) asked subjects to describe a word to another person. The authors manipulated the communicative intent of the speaker by claiming that the interlocutor either knew or did not know the

target word. In addition, Willems et al. (2010) manipulated the linguistic difficulty, by restricting the words a subject was allowed to use in the description. Specifically, the words had either a high or low semantic relation with the target word. The results indicated that a region in the mPFC was sensitive to the communicative demands of the task while the linguistic demands were represented in the left inferior frontal gyrus. Importantly, a region in the pSTS was sensitive to the interaction between linguistic and communicative demands of the task. A related study by Sassa and colleagues (2007) found that the mPFC, the temporal poles, and the left TPJ were sensitive to the communicative intention of the speaker. In summary, higher communicative task demands seem to correlate with stronger activation in areas within the classical ToM network.

In this study, IR sentences showed greater activation than control sentences in a set of regions, known to be involved in solving a specific ToM task, namely the false belief task (see Apperly, 2012, for discussion on the cognitive capacities underlying ToM). This was assessed using an ROI approach. During an independent scan, participants performed a classic ToM task (Dodell-Feder et al., 2010). Specifically, voxels were identified, which showed greater activation for interpreting false beliefs versus inaccurate physical descriptions (for a more detailed description, see Saxe & Kanwisher, 2003). As in previous studies, this contrast elicited more activation in a set of regions including the mPFC, the precuneus, and the bilateral TPJ. Some of these were also sensitive to the comprehension of IR sentences. In other words, IR sentences showed more activation in ToM areas than any of the three control conditions. This suggests that understanding IRs for action requires a similar inference on the mental state of the speaker as required for classical ToM tasks.

The results of this study show that the actual content of the statement does not trigger activation in motor areas. In the context of a picture of the desert, the sentence "It is hot here" does not result in motor activation. It is only in the case that this statement can be interpreted as a request for action that activation of motor areas is observed. It seems that the ToM network is needed to interpret the statement as a request in the right context. This suggests that the ToM system is involved in inferring what the speaker intends to convey with the string of words produced. In the case that the speaker intends to convey a request for action, areas involved in motor control get activated even when the utterance does not contain any lexical item that refers to an action. It is not easy to see how this could be accounted for in a Hebbian cell assembly account (Pulvermüller, 1999, 2005), which is based on associative connections between lexical items and actions. However, more recent accounts have acknowledged that different kinds of information might contribute to language understanding (Meteyard et al., 2012; Kiefer & Pulvermüller, 2012; Rueschemeyer, van Rooij, et al., 2010). In our case, the motor involve-

ment seems to presuppose the compositional machinery for decoding meaning and the inferential machinery for deriving speaker meaning in the situational context.

Summary and Conclusion

The current study investigated whether utterances with no explicit reference to an action activate the neural motor system if an action is communicatively implied. Specifically, brain responses to sentence–picture combinations, with IRs, were compared with control statements (PC, UC, and PUC). The results indicated that some parts of the cortical motor system were sensitive to both IR sentences (pre-SMA and bilateral IPL) and action execution. These findings extend previous research in showing that language explicitly referring to actions is not a necessary condition to elicit cortical motor activations. This suggests that motor areas might be involved in language comprehension in a much more complex way than theorists imagined just a decade ago (Pulvermüller, 1999, 2005). In addition, areas that were involved in thinking about mental states of others were also sensitive to IR sentences (mPFC, left TPJ). Very likely, these regions are crucial for making an inference about the communicative intent of the speaker.

Notes

1. Greenhouse–Geisser correction was used to correct for violations of sphericity.
2. The collection of stimuli was provided by Rebecca Saxe and Jessica Andrews-Hanna.

REFERENCES

Apperly, I. (2012). What is "theory of mind"? Concepts, cognitive processes and individual differences. *Quarterly Journal of Experimental Psychology, 65,* 825–839.

Aziz-Zadeh, L., & Damasio, A. (2008). Embodied semantics for actions: Findings from functional brain imaging. *Journal of Physiology Paris, 102,* 35–39.

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

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral Brain Science, 22,* 560–577.

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

Boulenger, V., Hauk, O., & Pulvermüller, F. (2011). Grasping ideas with the motor system: Semantic somatotopy in idiom comprehension. *Cerebral Cortex, 19,* 1905–1914.

Brett, M., Anton, J., Valabregue, R., & Poline, J. (2002). *Region of interest analysis using an SPM toolbox [abstract]*. 8th International Conference on Functional Mapping of the Human Brain, June 26, 2002, Sendai, Japan.

Caspers, S., Eickhoff, S. B., Geyer, S., Scheperjans, F., Mohlberg, H., Zilles, K., et al. (2008). The human inferior parietal lobule in stereotaxic space. *Brain Structure & Function, 212,* 481–495.

Caspers, S., Geyer, S., Schleicher, A., Mohlberg, H., Amunts, K., & Zilles, K. (2006). The human inferior parietal lobule: Cytoarchitectonic parcellation and interindividual variability. *Neuroimage, 33,* 430–448.

Desai, R. H., Binder, J. R., Conant, L. L., Mano, Q. R., & Seidenberg, M. S. (2011). The neural career of sensory-motor metaphors. *Journal of Cognitive Neuroscience, 23,* 2376–2386.

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2010). fMRI item analysis in a theory of mind task. *Neuroimage, 55,* 705–712.

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

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science, 308,* 662–667.

Frith, C., & Frith, U. (2005). Theory of mind. *Current Biology, 15,* 644–645.

Frith, U., & Frith, C. D. (2010). The social brain: Allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society, 365,* 165–176.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of "theory of mind." *Trends in Cognitive Sciences, 7,* 77–83.

Gallese, V., & Lakoff, G. (2005). The brain's concept's: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 21,* 455–479.

Geyer, S. (2004). The microstructural border between the motor and the cognitive domain in the human cerebral cortex. *Advances in Anatomy, Embryology and Cell Biology, 174,* 1–89.

Glenberg, A. M. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language, 43,* 379–401.

Glenberg, A. M., & Gallese, V. (2011). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex, 48,* 905–922.

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

Grice, H. P., Cole, P., & Morgan, J. (1975). Logic and conversation. In P. Cole & J. L. Morgen (Eds.), *Syntax and semantics, Vol. 3, Speech acts* (pp. 45–47). New York: Academic Press.

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences, 9,* 416–423.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41,* 301–307.

Holtgraves, T. M. (1994). Communication in context: Effects of speaker status on the comprehension of indirect requests. *Journal of Experimental Psychology: Learning Memory and Cognition, 20,* 1205–1218.

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

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought.* New York: Basic Books.

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, 102,* 59–70.

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

Picard, N., & Strick, P. L. (2001). Imaging the premotor areas. *Current Opinion in Neurobiology, 11,* 663–672.

Postle, N., McMahon, K., Ashton, R., & de Zubicaray, G. I. (2008). Action word meaning representations in cytoarchitectonically defined primary and premotor cortices. *Neuroimage, 43,* 634–644.

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

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

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

Rueschemeyer, S. A., van Rooij, D., Lindemann, O., Willems, R. M., & Bekkering, H. (2010). The function of words: Distinct neural correlates for words denoting differently manipulable objects. *Journal of Cognitive Neuroscience, 22,* 1844–1851.

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

Rushworth, M. F. S., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences, 8,* 410–417.

Sassa, Y., Sugiura, M., Jeong, H., Horie, K., Sato, S., & Kawashima, R. (2007). Cortical mechanisms of communicative speech production. *Neuroimage, 37,* 985–992.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people—The role of the temporo-parietal junction in "theory of mind." *Neuroimage, 19,* 1835–1842.

Slotnick, S. D., Moo, L. R., Segal, J. B., & Hart, J. (2003). Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Brain Research, 17,* 75–82.

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.

Van Dam, W., Rueschemeyer, S.-A., & Bekkering, H. (2010). Action specificity reflected in embodied lexical-semantic representations. *Neuroimage, 53,* 1318–1325.

Van Dam, W. O., van Dijk, M., Bekkering, H., & Rueschemeyer, S.-A. (2011). Flexibility in embodied lexical-semantic representations. *Human Brain Mapping.* doi: 10.1002/hbm.21365.

Wheaton, L. A., & Hallett, M. (2007). Ideomotor apraxia: A review. *Journal of the Neurological Sciences, 260,* 1–10.

Willems, R. M., de Boer, M., de Ruiter, J. P., Noordzij, M. L., Hagoort, P., & Toni, I. (2010). A dissociation between linguistic and communicative abilities in the human brain. *Psychological Science, 21,* 8–14.

Willems, R. M., Ludovica, 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.