Brain Function Overlaps When People Observe Emblems, Speech, and Grasping

Michael Andric[a], **Ana Solodkin**[b], **Giovanni Buccino**[d], **Susan Goldin-Meadow**[a], **Giacomo Rizzolatti**[e], and **Steven L. Small**[c]

[a]Department of Psychology, The University of Chicago, Chicago, IL, USA;

[b]Departments of Anatomy & Neurobiology and Neurology. University of California, Irvine School of Medicine, Irvine, CA, USA

[c]Department of Neurology, University of California, Irvine School of Medicine, Irvine, CA, USA

[d]Department of Medical and Surgical Sciences, University Magna Graecia, Catanzaro, Italy; Istituto Neurologico Mediterraneo Neuromed, Pozzilli (Is), Italy

[e]Dipartimento di Neuroscienze, Università di Parma, Parma, Italy

Abstract

A hand grasping a cup or gesturing 'thumbs-up', while both manual actions, have different purposes and effects. Grasping directly affects the cup, whereas gesturing 'thumbs-up' has an effect through an implied verbal (symbolic) meaning. Because grasping and emblematic gestures ('emblems') are both goal-oriented hand actions, we pursued the hypothesis that observing each should evoke similar activity in neural regions implicated in processing goal-oriented hand actions. However, because emblems express symbolic meaning, observing them should also evoke activity in regions implicated in interpreting meaning, which is most commonly expressed in language. Using fMRI to test this hypothesis, we had participants watch videos of an actor performing emblems, speaking utterances matched in meaning to the emblems, and grasping objects. Our results show that lateral temporal and inferior frontal regions respond to symbolic meaning, even when it is expressed by a single hand action. In particular, we found that left inferior frontal and right lateral temporal regions are strongly engaged when people observe either emblems or speech. In contrast, we also replicate and extend previous work that implicates parietal and premotor responses in observing goal-oriented hand actions. For hand actions, we found that bilateral parietal and premotor regions are strongly engaged when people observe either emblems or grasping. These findings thus characterize converging brain responses to shared features (e.g., symbolic or manual), despite their encoding and presentation in different stimulus modalities.

Keywords

gestures; language; semantics; perception; functional magnetic resonance imaging

---

Corresponding Author: Michael Andric Center for Mind/Brain Sciences (CIMeC) The University of Trento Via delle Regole 101 38060 Mattarello (TN), Italy michael.andric@unitn.it Phone: +39 0461 283660.

1. INTRODUCTION

People regularly use their hands to communicate, whether to perform gestures that accompany speech ('co-speech gestures') or to perform gestures that – on their own – communicate specific meanings, e.g., performing a "thumbs-up" to express "it's good." These latter gestures are called "emblematic gestures" – or "emblems", and require a person to process both the action and its implied verbal (symbolic) meaning. Action observation and meaning processing are highly active areas of human neuroscience research, and significant research has examined the way that the brain processes meaning conveyed with the hands. Most of this research has focused on conventional sign language and co-speech gestures, not on emblems. Emblems differ from these other types of gesture in fundamental ways. Although individual emblems express symbolic meaning, they do not use the linguistic and combinatorial structures of sign language, which is a fully developed language system. Emblems also differ from co-speech gestures, which require accompanying speech for their meaning (McNeill, 2005). Thus, in contrast with sign language, emblems are not combinatorial and lack the linguistic structures found in human language. In contrast with co-speech gestures, emblems can directly convey meaning in the absence of speech (Ekman & Friesen, 1969; Goldin-Meadow, 1999, 2003; McNeill, 2005).

At the same time, emblems are manual actions, and as such, are visually similar to actions that are not communicative, such as manual grasping. Emblems also represent a fundamentally different way of communicating symbolic meaning compared to spoken language. Although the lips, tongue, and mouth perform actions during speech production, these movements per se neither represent nor inform the meaning of the utterance. Thus, from the biological standpoint, the brain must encode and operate on emblems in two ways, (i) as meaningful symbolic expressions, and (ii) as purposeful hand actions (Figure 1). The ways that these two functions are encoded, integrated, and applied in understanding emblems is the subject of the present study.

Processing symbolic meaning expressed in language engages many disparate brain areas, depending on the type of language used and the goal of the communication. But some brain areas are highly replicated across these diverse communicative contexts. For example, a recent meta-analysis described semantic processing to primarily involve parts of the lateral and ventral temporal cortex, left inferior frontal gyrus, left middle and superior frontal gyri, left ventromedial prefrontal cortex, the supramarginal (SMG) and angular gyri (AG), and the posterior cingulate cortex (Binder, Desai, Graves, & Conant, 2009). More specifically, posterior middle temporal gyrus (MTGp) responses have often been associated with recognizing word meanings (Binder, et al., 1997; Chao, Haxby, & Martin, 1999; Gold, et al., 2006), and anterior superior temporal activity has been associated with processing combinations of words, such as phrases and sentences (Friederici, Meyer, & von Cramon, 2000; Humphries, Binder, Medler, & Liebenthal, 2006; Noppeney & Price, 2004). In the inferior frontal gyrus (IFG), pars triangularis (IFGTr) activity has often been found when people discriminate semantic meaning (Binder, et al., 1997; Devlin, Matthews, & Rushworth, 2003; Friederici, Opitz, & Cramon, 2000), while pars opercularis (IFGOp) function has been linked with a number of tasks. Some of these tasks involve audiovisual speech perception (Broca, 1861; Hasson, Skipper, Nusbaum, & Small, 2007; Miller & D'Esposito, 2005), but others involve recognizing hand actions (Binkofski & Buccino, 2004; Rizzolatti & Craighero, 2004).

Prior biological work on the understanding of observed hand actions implicates parietal and premotor cortices. In the macaque, parts of these regions interact to form a putative "mirror system" that is thought to be integral in action observation and execution (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Fogassi, Gallese, Fadiga, & Rizzolatti, 1998;

Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). A similar system appears to be present in humans, and also to mediate human action understanding (Fabbri-Destro & Rizzolatti, 2008; Rizzolatti & Arbib, 1998; Rizzolatti & Craighero, 2004; Rizzolatti, Fogassi, & Gallese, 2001). Studies investigating human action understanding have, in fact, found activity in a variety of parietal and premotor regions when people observe hand actions. This includes object-directed actions, such as grasping (Buccino, et al., 2001; Grezes, Armony, Rowe, & Passingham, 2003; Shmuelof & Zohary, 2005, 2006), and non-object-directed actions, such as pantomimes (Buccino, et al., 2001; Decety, et al., 1997; Grezes, et al., 2003). More precisely, some of the parietal regions involved in these circuits include the intraparietal sulcus (IPS) (Buccino, et al., 2001; Buccino, et al., 2004; Grezes, et al., 2003; Shmuelof & Zohary, 2005, 2006) and inferior and superior parietal lobules (Buccino, et al., 2004; Perani, et al., 2001; Shmuelof & Zohary, 2005, 2006). In the premotor cortex, this includes the ventral (PMv) and dorsal (PMd) segments (Buccino, et al., 2001; Grezes, et al., 2003; Shmuelof & Zohary, 2005, 2006). Because emblems are hand actions, perceiving them should also involve responses in these areas. However, it remains an open question the extent to which these areas are involved in emblem processing. Further, the anatomical and physiological mechanisms used by the brain to decode the integrated manual and symbolic features of emblematic gestures are not known.

Recently, an increasing number of studies have sought to understand the way that the brain gleans meaning from manual gestures, particularly co-speech gestures. In general, co-speech gestures appear to activate parietal and premotor regions (Kircher, et al., 2009; J. I. Skipper, Goldin-Meadow, Nusbaum, & Small, 2009; Villarreal, et al., 2008; Willems, Ozyurek, & Hagoort, 2007). Yet, activity during co-speech gesture processing has also been found in regions associated with symbolic meaning (see Binder, et al., 2009 for review). These regions include parts of the IFG, such as the IFGTr (Dick, Goldin-Meadow, Hasson, Skipper, & Small, 2009; J. I. Skipper, et al., 2009; Willems, et al., 2007) and lateral temporal areas, such as the MTGp (Green, et al., 2009; Kircher, et al., 2009).

It is not surprising that areas that respond when people comprehend language also respond when people comprehend gestures in the presence of spoken language. Several studies thus attempt to disentangle the brain responses specific to the meaning of co-speech gestures from those of the accompanying language. Typically, this is done by contrasting audiovisual speech containing gestures with audiovisual speech without gestures (Green, et al., 2009; Willems, et al., 2007). By way of subtractive analyses, the results generally reflect greater activity in these 'language' areas when gestures accompany speech than when they don't. Greater activity in these areas is then taken as a measure of their importance in determining meaning (J. I. Skipper, et al., 2009; Willems, et al., 2007).

However, co-speech gestures are processed interactively with accompanying speech (Bernardis & Gentilucci, 2006; Gentilucci, Bernardis, Crisi, & Dalla Volta, 2006; Kelly, Barr, Church, & Lynch, 1999), and it is the accompanying speech that gives co-speech gestures their meaning (McNeill, 2005). In other words, speech and gesture information do not simply add up in a linear way. Thus, when the hands express symbolic information, it is difficult to truly separate the brain responses attributable to gestural meaning from those of the accompanying spoken language.

Previous research to examine brain responses to emblems does not present a clear profile of activity that characterizes how the brain comprehends them. This may be due partly to the wide variation in methods and task demands in these studies. Indeed, prior emblem research has been tailored to address such diverse questions as their social relevance (Knutson, McClellan, & Grafman, 2008; Lotze, et al., 2006; Montgomery, Isenberg, & Haxby, 2007;

Straube, Green, Jansen, Chatterjee, & Kircher, 2010), emotional salience (Knutson, et al., 2008; Lotze, et al., 2006), or shared symbolic basis with pantomimes and speech (Xu, Gannon, Emmorey, Smith, & Braun, 2009). Accordingly, the results implicate a disparate range of brain areas. These areas include the left IFG (Lindenberg, Uhlig, Scherfeld, Schlaug, & Seitz, 2012; Xu, et al., 2009), right IFG (Lindenberg, et al., 2012; Villarreal, et al., 2008), insula (Montgomery, et al., 2007), premotor cortex (Lindenberg, et al., 2012; Montgomery, et al., 2007; Villarreal, et al., 2008), MTG (Lindenberg, et al., 2012; Villarreal, et al., 2008; Xu, et al., 2009), right (Xu, et al., 2009) and bilateral fusiform gyri (Villarreal, et al., 2008), left (Lotze, et al., 2006) and bilateral inferior parietal lobules (Montgomery, et al., 2007; Villarreal, et al., 2008), medial prefrontal cortex (Lotze, et al., 2006; Montgomery, et al., 2007), as well as the temporal poles (Lotze, et al., 2006; Montgomery, et al., 2007). This represents a very large set of brain responses to emblems and does not clarify the question of interest here, namely the mechanisms underlying the decoding of symbolic and manual information.

In the present study, we aimed 1) to identify brain areas that decode symbolic meaning, independent of its expression as emblem or speech, and 2) to identify brain areas that process hand actions, regardless of whether they are symbolic emblems or non-symbolic grasping actions. To identify brain areas sensitive to symbolic meaning, we had participants watch an actor communicate similar meanings with speech (e.g., saying "It's good") and with emblems (e.g., performing a "thumbs-up" to symbolize "It's good"; see Figure 2). With this experimental manipulation, we sought to identify their common neural basis as expressions of symbolic meaning. Thus we were not interested in the differences between emblems and speech, but in their similarities. In other words, despite emblems and speech having many differentiating perceptual and/or cognitive features, their shared responses could represent those areas sensitive to perceiving symbolic meaning in both. Similarly, to characterize brain regions associated with hand action processing, participants also saw the actor perform object-directed grasping. As with speech, we were not focused on the differences between emblems and grasping, but on their similarities. This allowed us to identify brain areas active during hand action observation, both in the context of symbolic expression and to non-symbolic actions, such as grasping an object.

Synthesizing previous findings on how the brain processes meaning conveyed in language and by co-speech gestures, we expected that interpreting symbolic meaning, independent of its mode of presentation (as speech or emblems), would largely associate with overlapping anterior inferior frontal, MTGp, and anterior superior temporal gyrus (STGa) activity. Conversely, we expected that observing hand actions – both those performed with or without an object – would lead to responses in such areas as the IPS, inferior and superior parietal lobules, as well as the ventral and dorsal premotor cortices. In summary, we postulated that activity in one set of regions would converge when perceiving symbolic meaning, and in another set for perceiving hand actions, independent of their symbolic (emblems, speech) or object-directed (emblems, grasping) basis.

2. MATERIALS AND METHODS

2.1. Participants

Twenty-four people (14 women, mean age 21.13 years, SD = 2.94) were recruited from the student population of The University of Chicago. All were right-handed (score mean = 82.57, SD = 15.46, range = 50–100; Oldfield, 1971), except for one that was slightly ambidextrous (score = 20). All participants were native speakers of American English with normal hearing and vision and no reported history of neurological or psychological disturbance. The Institutional Review Board of The University of Chicago approved the study, and all participants gave written informed consent.

2.2. Stimuli

Our stimuli consisted of 3-4 second long video clips in three experimental conditions. One condition (Emblem) showed a male actor performing emblematic gestures (e.g., giving a thumbs-up, a pinch of the thumb and index finger to form a circle with other fingers extended, a flat palm facing the observer, a shrug of the shoulders with both hands raised). In a second condition (Speech), the actor said short phrases that expressed meanings similar to the meanings conveyed by the emblems (e.g., "It's good", "Okay", "Stop", "I don't know"). The emblems were chosen so that they referred to the meanings of the words in the Speech condition. For example, "It's good" was matched with the emblem "raise the arm with a closed fist and the thumb up". In the third condition (Grasping), the actor was shown grasping common objects (e.g., a stapler, a pen, an apple, a cup). There were 48 videos per condition and 192 videos total. This includes 48 videos of a fourth condition for which data was collected during the experiment. However, this condition was intended for a separate investigation. It is not analyzed here and not discussed further.

A male native speaker of American English was used as the actor in the same setting for all videos. The actor was assessed as strongly left handed by the Edinburgh handedness inventory, so all hand actions were performed with the left hand, his dominant and preferred manual effector. The actor made no noticeable facial movements besides those used in articulation and directed his gaze in ways that were naturally congruent with the performed action (toward the audience in the emblem and speech videos and toward the object when grasping). To generate a set of "right handed" actions, the original videos were horizontally flipped. To check the ecological validity of these flipped videos, they were shown to a set of individuals who were blind to the experimental protocol. These individuals were asked to look for anything unusual or unexpected in the clips; no abnormalities were reported. Experimental items were chosen from an initial set (n = 35) by selecting only videos that elicited the same meanings in a separate sample at The University of Chicago (n = 10).

2.3. Procedure

In the scanner, to determine a comfortable sound level for each participant, we played a practice clip while the scanner emitted sounds heard during a functional scan. Following this sound level calibration, participants passively viewed the video clips in 6 separate runs. Each run was 330 s long. Natural viewing allowed avoiding systematic bias in participants' gaze that might otherwise mask the effects of interest (Wang, Ramsey, & de, 2011). To avoid ancillary task requirements, no explicit responses were required of the participants in the scanner. Half of the participants viewed actions performed with the left hand (LV). The other half viewed right-handed actions (RV). A stimulus onset asynchrony of 20 s was used after an initial 10 s of rest at the beginning of each run in a slow event-related design. During the initial 10 s of rest, as well as during the period between the end of each clip and the onset of the next, participants saw an empty black screen. The participants heard audio through headphones. The videos were viewed via a mirror that was attached to the head coil, allowing participants to see a screen at the end of the scanning bed.

2.4. Image Acquisition and Data Analyses

Scans were acquired at 3 Tesla using spiral acquisition (Noll, Cohen, Meyer, & Schneider, 1995) with a standard head coil. For each participant, two volumetric T1-weighted scans (120 axial slices, $1.5 \times 0.938 \times 0.938$ mm resolution) were acquired and averaged. This provided high-resolution images on which anatomical landmarks could be identified and functional activity maps could be overlaid. Functional images were collected across the whole brain in the axial plane with TR = 2000 ms, TE = 30 ms, FA = 77 degrees, in 32 slices with thickness of 3.8 mm for a voxel resolution of $3.8 \times 3.75 \times 3.75$ mm. The images were registered in 3D space by Fourier transformation of each of the time points and

corrected for head movement using AFNI (Cox, 1996). The time series data was mean normalized to percent signal change values. Then, the hemodynamic response function (HRF) for each condition was established via a regression for the 18 s following the stimulus presentation on a voxel-wise basis. There were separate regressors in the model for each of the four experimental conditions. Additional regressors were the mean, linear, and quadratic trend components, as well as the 6 motion parameters in each of the functional runs. A linear least squares model was used to establish a fit to each time point of the HRF for each of the four conditions.

We used FreeSurfer (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999) to create surface representations of each participant's anatomy. Each hemisphere of the anatomical volumes was inflated to a surface representation and aligned to a template of average curvature. The functional data were then projected from the 3D volumes onto the 2-dimensional surfaces using SUMA (Saad, 2004). Doing this enables more accurate reflection of the individual data at the group level (Argall, Saad, & Beauchamp, 2006). To decrease spatial noise, the data were then smoothed on the surface with a Gaussian 4-mm FWHM filter. These smoothed values for each participant were next brought into a MySQL relational database. This allowed the data to be queried in statistical analyses using R (http://www.r-project.org).

2.4.1 Whole-brain analyses—To identify the brain's task-related activation (signal change) with respect to a resting baseline for each condition, we performed two vertex-wise analyses across the cortical surface. The reliability of clusters was determined using a permutation approach (Nichols & Holmes, 2002), which identified significant clusters with an individual vertex threshold of $p < .001$, corrected for multiple comparisons to achieve a family-wise error (FWE) rate of $p < .05$. Clustering proceeded separately for positive and negative values. The first analysis investigated any between-group differences for the LV and RV groups. Comparisons were specified for each of the experimental conditions to assess any reliable differences for observing the actions performed with either the left or right hand (individual vertex threshold, $p < .01$, FWE $p < .05$). Only one reliable cluster of activity was found in this analysis: for the Emblem condition, we found LV > RV in the inferior portion of the left post-central sulcus. Thus, we performed further analyses, collapsing across the LV and RV groups to include all 24 participants. Brain areas sensitive to observing the experimental stimuli involving hand actions were found by examining the intersection ("conjunction", Nichols, Brett, Andersson, Wager, & Poline, 2005) of brain activity from the direct whole-brain contrasts. This analysis identified overlapping activity in the Emblem and Grasping conditions, both significantly active above a resting baseline. This yielded a map of Emblem & Grasping. Similarly, to assess areas across the brain that showed sensitivity for perceiving symbolic meaning, independent of its presentation as emblem or speech, we examined the conjunction of significant activity in the Emblem and Speech conditions (Emblem & Speech). Conjunction maps of Grasping & Speech, and for all three conditions, were also generated for comparative purposes. Finally, though outside the experimental questions in this paper, an additional vertex-wise analysis that compared above baseline activity between conditions was done. These exploratory findings were also determined using the cluster thresholding procedure described above. They are presented as supplemental material.

2.4.2. Region of Interest Analysis—To further evaluate regions in which activity from the cortical surface analysis was significant, we examined activity (signal change) in anatomically defined regions of interest (ROIs). The regions were delineated on each individual's cortical surface representation, using an automated parcellation scheme (Desikan, et al., 2006; Fischl, et al., 2004). This procedure uses a probabilistic labeling algorithm that incorporates the anatomical conventions of Duvernoy (Duvernoy, 1991) and

has a high accuracy approaching that of manual parcellation (Desikan, et al., 2006; Fischl, et al., 2002; Fischl, et al., 2004). We manually augmented the parcellation with further subdivisions: superior temporal gyrus (STG) and superior temporal sulcus (STS) were each divided into anterior and posterior segments; and the precentral gyrus was divided into inferior and superior parts. The following regions were tested: pars opercularis (IFGOp), pars triangularis/orbitalis (IFGTr/IFGOr), ventral premotor (PMv), dorsal premotor (PMd), anterior superior temporal gyrus (STGa), posterior superior temporal sulcus (STSp), posterior middle temporal gyrus (MTGp), supramarginal gyrus (SMG), intraparietal sulcus (IPS), and superior parietal (SP).

Data analysis was carried out separately for each region in each hemisphere. The dependent variable for this analysis was the percent signal change in the area under the hemodynamic curve for time points 2 thru 6 (seconds 2 – 10). This area comprised 75% of the HRF in the calcarine fissure, expected to include primary visual cortex. We selected these time points in order to isolate the dominant component of the HRF in all regions of the brain. We selected the calcarine fissure as one exemplar because every condition included visual information. In addition, we validated selection of these time points by examining the HRF in transverse temporal gyrus (TTG), a region that includes primary auditory cortex, for the speech condition. We filtered vertices that contributed outlying values in the region by normalizing the percent signal change value for each vertex and removing those that were greater than 2.5 SDs away from the mean of the region for that participant. In order to gain information about differences between conditions where activity was above baseline, the data were thresholded at each vertex in the region to include only positive activation. Significant differences between conditions in the regions were assessed using paired t-tests.

3. RESULTS

We present the results of three main analyses. First, we identified brain areas associated with observing emblems, grasping, and speech. Second, we examined the convergence of activity shared by emblems and speech (symbolic meaning) and by emblems and grasping (symbolic and non-symbolic hand actions). Finally, to further assess the involvement of specific regions of interest (ROIs) thought to be involved in either symbolic or hand action-related encodings, we examined signal intensities across conditions.

3.1. General activity for observing emblems, grasping, and speech

Figure 3 shows brain activity for the Emblem, Grasping, and Speech conditions (per vertex, $p < .001$, FWE $p < .05$). Among areas active in all conditions, we found activity in the posterior superior temporal sulcus (STSp), the MTGp, as well as primary and secondary visual areas. These visual areas included the middle occipital gyrus and anterior occipital sulcus. Grasping observation elicited bilateral activity in parietal areas, such as the IPS and SMG, as well as the PMv and PMd cortex. Activation for observing speech included bilateral transverse temporal gyrus (TTG), STGa, posterior superior temporal gyrus (STGp), MTGp, the IFGTr and pars orbitalis (IFGOr), the IFGOp, and the left SMG. Further activation for observing emblems extended across widespread brain areas. These areas comprised much of the parietal and premotor cortices, as well as inferior frontal and lateral temporal areas.

3.2. Converging activity: Symbolic meaning

The intersection of statistically significant activity between the speech and the emblem conditions showed bilateral STS and visual cortices, as well as lateral temporal and frontal cortices (Figure 4A, yellow). Specifically, bilateral MTGp activity spread through the STS and along the STG, extending in the right hemisphere to the STGa. Convergent areas in

frontal cortex covered not only bilateral IFG, including parts of IFGTr and IFGOp, but also bilateral PMv and PMd.

3.3. Converging activity: Hand actions

Areas demonstrating convergence related to perceived hand actions were predominantly found in parietal areas, specifically the IPS and SMG, bilaterally, and frontal areas, including PMv and PMd, as well as the IFG (Figure 4B, yellow). Noticeably different from those converging areas for symbolic meaning is that much of the activity for symbolic meaning spread anterior both along the STG and in the left IFG. Furthermore, in frontal cortex, we found convergent premotor activity between emblems and grasping in a section immediately posterior to that for emblems and speech.

3.4. ROI Analysis

3.4.1. Symbolic regions stronger than hand regions—In a number of lateral temporal and inferior frontal regions, we found stronger Emblem and Speech activity compared to Grasping – but not compared to each other (Figure 5). Specifically, these regions included the right MTGp (both comparisons, $p < .01$) and STGa (Emblem > Grasping, $p < .05$ and Speech > Grasping, $p < .001$). This was also true in the inferior frontal cortex, specifically the left IFGTr/IFGOr (Emblem > Grasping, $P < .01$ and Speech > Grasping, $p < .001$). In the left IFGOp, Emblem and Speech activity was also stronger compared to Grasping. But this was statistically significant ($P < .001$) only for Speech compared to Grasping.

3.4.2. Hand regions stronger than symbolic regions—In contrast, premotor and parietal regions responded stronger when people observed hand actions, i.e., emblems and grasping, as opposed to speech (Figure 6). Specifically, in PMv and PMd, bilaterally, activity was significantly stronger for both Emblem and Grasping compared to Speech. Activity did not significantly differ between the manual conditions though (Figure 6, right side). Similarly, compared to Speech, both Emblem and Grasping elicited stronger activity in the IPS and superior parietal cortices, bilaterally (Figure 6, left side).

3.4.3. Lateralized SMG responses—In contrast with regions listed above, the SMG did not uniformly respond across hemispheres to either hand actions or symbolic meaning. Rather, SMG activity differed between hemispheres (Figure 7). The left SMG responded significantly stronger for Speech compared to Grasping ($p = .001$). Conversely, the right SMG responded significantly stronger for Grasping compared to Speech ($p < .05$). We found an intermediate response for Emblem in both the left and right SMG, where Emblem activity did not significantly differ from either Grasping or Speech.

4. DISCUSSION

Our results show that when people observe emblems the brain produces activity that substantially overlaps with activity produced while observing both speech and grasping. For processing meaning, these overlapping responses assert the importance of lateral temporal and inferior frontal regions. For processing hand actions, our findings replicate previous work that highlights the importance of parietal and premotor responses. Specifically, the right MTGp and STGa, as well as the left IFGTr/IFGOr, are active in processing meaning – regardless of whether it is conveyed by emblems or speech. These lateral temporal and inferior frontal responses are also stronger for processing symbolic meaning compared to non-symbolic grasping actions. In contrast, regions such as the IPS, superior parietal cortices, PMv, and PMd respond to hand actions – regardless of whether the actions are symbolic or object-directed. Activity in these parietal and premotor regions is also

significantly stronger for hand actions than for speech. Thus, emblem processing incorporates visual, parietal, and premotor responses that are often found in action observation with inferior frontal and lateral temporal responses that are common in language understanding. This suggests that brain responses may be organized at one level by perceptual recognition (e.g., visually perceiving a hand) but at another by the type of information to be interpreted (e.g., symbolic meaning).

4.1. Processing symbolic meaning

We found that when people perceived either speech or emblems the right MTGp and STGa, as well as the left IFG, significantly respond. These regions' convergence implicates their sensitivity beyond perceptual encoding, to the level of processing meaning – regardless of its codified form (i.e., spoken or manual). This implication agrees with previous findings, both for verbally communicated meaning and manual gesture.

However, it is worth reiterating that the different gesture types used in previous studies, which identify temporal and inferior frontal activity, vary in the ways they convey meaning. Whereas co-speech gestures, by definition, use accompanying speech to convey meaning, emblems can do so on their own. In addition, gestures' varying social content can play a role in modulating responses (Knutson, et al., 2008). For example, brain function differs as a function of the communicative mean employed to convey a communicative intention perceived by an observer (Enrici, Adenzato, Cappa, Bara, & Tettamanti, 2011). It can also vary when people are addressed directly versus indirectly, i.e., depending on whether an actor directly faces the observer (Straube, et al., 2010).

Indeed, the MTGp's association with processing meaning persists across numerous contexts that involve language and gestures. For example, significant MTGp activity has repeatedly been found in word recognition (Binder, et al., 1997; Fiebach, Friederici, Muller, & von Cramon, 2002; Mechelli, Gorno-Tempini, & Price, 2003) and generation (Fiez, Raichle, Balota, Tallal, & Petersen, 1996; Martin & Chao, 2001). MTGp activity has also been found using lexical-semantic selection tasks (Gold, et al., 2006; Indefrey & Levelt, 2004).

MTGp activity has often been found in gesture processing, as well. This includes co-speech gestures (Green, et al., 2009; Kircher, et al., 2009; Straube, Green, Bromberger, & Kircher, 2011; Straube, et al., 2010), iconic gestures without speech (Straube, Green, Weis, & Kircher, 2012), emblems (Lindenberg, et al., 2012; Villarreal, et al., 2008; Xu, et al., 2009), and pantomimes (Villarreal, et al., 2008; Xu, et al., 2009). In addition, lesion studies have suggested this region's function is critically important when people identify an action's meaning (Kalenine, Buxbaum, & Coslett, 2010).

Showing significant responses to these multiple stimulus types, MTGp function then does not appear to be modality specific. Instead, its sensitivity seems more general. That is, its responses are evidently not tied to just verbal or gesture input per se. A recent study, in fact, found MTGp activity both when people perceive spoken sentences as well as iconic gestures without speech (Straube, et al., 2012). While this region was classically suggested to be part of visual association cortex (Mesulam, 1985; von Bonin & Bailey, 1947), its function in auditory processing is also well-documented (Hickok & Poeppel, 2007; Humphries, et al., 2006; Wise, et al., 2000; Zatorre, Evans, Meyer, & Gjedde, 1992). Moreover, the MTGp's properties as heteromodal cortex have led some authors to suggest it as important for "supramodal integration and conceptual retrieval" (Binder, et al., 2009). Our results agree with this interpretation. With MTGp activity found here both in response to speech and emblems, this region's importance in higher-level functions, such as conceptual processing – without modality dependence – appears likely.

Similarly, we also found significant STGa activity for processing meaning conveyed by either emblems or speech. STGa activity has been associated with interpreting verbally communicated meaning. For example, activity in this region has been found when people process sentences (Friederici, Meyer, et al., 2000; Humphries, Love, Swinney, & Hickok, 2005; Noppeney & Price, 2004), build phrases (Brennan, et al., 2010; Humphries, et al., 2006), and determine semantic coherence (Rogalsky & Hickok, 2009; Stowe, Haverkort, & Zwarts, 2005). In the current experiment, verbal information was presented only in speech (e.g., "It's good", "Stop", "I don't know"). But both speech and emblems conveyed coherent semantic information.

Our finding that the STGa responds to both speech and emblems thus extends its role. In other words, the presence of STGa activity when people process either speech or emblems suggests that this region's responses are not based simply on verbal input. Rather, the STGa appears more generally tuned for perceiving coherent meaning across multiple forms of representation. In fact, a recent review of anterior temporal cortex function suggests it acts as a semantic hub (Patterson, Nestor, & Rogers, 2007). From this perspective, semantic information may be coded beyond the stimulus features used to convey it. For example, visually perceiving gestures would evoke visuo-motor responses. But, apart from the visual and motor features used to convey it, semantic content conveyed by a gesture would further involve anterior temporal responses that are particularly tuned for this type of information. Indeed, our results agree with this account and corroborate this area's suggested amodal sensitivity (Patterson, et al., 2007).

We also implicated the left IFG in processing meaning for both speech and emblems. For the anterior IFG (i.e., IFGTr/IFGOr), this was already known. For example, IFGTr activity has been found when people determine meaning, both from language (Dapretto & Bookheimer, 1999; Devlin, et al., 2003; Friederici, Opitz, et al., 2000; Gold, Balota, Kirchhoff, & Buckner, 2005; Wagner, Pare-Blagoev, Clark, & Poldrack, 2001) and gestures (Kircher, et al., 2009; Molnar-Szakacs, Iacoboni, Koski, & Mazziotta, 2005; Jeremy I. Skipper, Goldin-Meadow, Nusbaum, & Small, 2007; Straube, et al., 2011; Villarreal, et al., 2008; Willems, et al., 2007; Xu, et al., 2009). The importance of this region in processing represented meaning has also been documented in patient studies. For example, left frontoparietal lesions that include the IFG have been associated with impaired action recognition – even when the action has to be recognized through sounds typically associated with the action (Pazzaglia, Pizzamiglio, Pes, & Aglioti, 2008; Pazzaglia, Smania, Corato, & Aglioti, 2008).

In contrast, left IFGOp activity has been linked with a wide range of language and motor processes. For example, the left IFGOp has been associated with perceiving audiovisual speech (Broca, 1861; Hasson, et al., 2007; Miller & D'Esposito, 2005) and when people interpret co-speech gestures (Green, et al., 2009; Kircher, et al., 2009). Similarly, the IFGOp might also be important for recognizing mouth and hand actions, respectively, in the absence of language or communication (see Binkofski & Buccino, 2004; Rizzolatti & Craighero, 2004). Our results show this area to be active in all conditions. Yet, the strongest responses were for perceiving speech and emblems (more than grasping). Thus, left IFGOp responses may be more preferentially tuned to respond to mouth and hand actions that convey symbolic meaning.

4.2. Processing hand actions

We found overlapping parietal and premotor responses when people observed either emblems or manual grasping. Specifically, bilateral PMv and PMd activity, as well as IPS and superior parietal lobe activity, were elicited during the manual conditions. Also, responses in these regions were stronger for the manual action conditions than for speech.

Our findings in these regions are consistent with prior data on observing manual grasping (Grezes, et al., 2003; Manthey, Schubotz, & von Cramon, 2003; Shmuelof & Zohary, 2005, 2006) and gestures (Enrici, et al., 2011; Hubbard, Wilson, Callan, & Dapretto, 2009; Lui, et al., 2008; Jeremy I. Skipper, et al., 2007; Straube, et al., 2012; Villarreal, et al., 2008; Willems, et al., 2007). Our findings also agree with an extensive patient literature that links parietal and premotor damage to limb apraxias (see Leiguarda & Marsden, 2000 for review). By demonstrating that these regions are similarly involved in perceiving emblems as in perceiving grasping, we further generalize their importance in comprehending hand actions. That is, our results implicate parietal and premotor function more generally, at the level of perceiving hand actions, rather than differentiating their particular uses or goals.

Recent findings from experiments with macaque monkeys implicate parietal and premotor cortices in understanding goal-directed hand actions (Fabbri-Destro & Rizzolatti, 2008; Rizzolatti & Craighero, 2004; Rizzolatti, et al., 2001). For example, it has been found that neurons in macaque ventral premotor cortex area F5 (di Pellegrino, et al., 1992; Rizzolatti, et al., 1988) and inferior parietal area PF (Fogassi, et al., 1998) code specific actions. That is, there are neurons in these areas that fire both when the monkey performs an action and when it observes another performing the same or similar action.

Some recent research has tried to identify whether there are homologous areas in humans that code specific actions (e.g., Grezes, et al., 2003). As noted, many human studies have, in fact, characterized bilateral parietal and premotor responses during grasping observation. Activity in these regions has also been associated with viewing non-object-directed actions, such as pantomimes (Buccino, et al., 2001; Decety, et al., 1997; Grezes, et al., 2003). However, pantomimes characteristically require an object's use without its physical presence. Thus, it is ambiguous whether pantomimes are truly 'non-object-directed' actions (see Bartolo, Cubelli, Della Sala, & Drei, 2003 for discussion). Prior to the present study, it was not clear whether communicative, symbolic actions (i.e., emblems) and object-directed, non-symbolic actions evoked similar brain responses.

Do communicative symbolic actions elicit similar responses in parietal and premotor areas as object-directed actions? This was an outstanding question that lead to the current investigation. Previous studies of symbolic gestures, including co-speech gestures and emblems, have mostly focused on responses to specific features of these actions (e.g., their iconic meaning or social relevance). Some of these differing features likely evoke brain responses that diverge from responses to other manual actions, such as grasping. For example, responses in medial frontal areas, associated with processing others' intentions (Mason, et al., 2007) and mental states (Mason, Banfield, & Macrae, 2004), have been reported for emblem processing (Enrici, et al., 2011; Straube, et al., 2010). Such responses are not typically associated with grasping observation. Also, as described above, emblems' communicative effect appears to evoke responses that are shared for processing speech, again contrasting with grasping.

Still, hand actions, despite their varying characteristics, may share a common neural basis, more generally. For example, significant parietal and premotor responses are reported in some previous co-speech gesture studies (Kircher, et al., 2009; J. I. Skipper, et al., 2009; Willems, et al., 2007), as well as some studies of emblems (e.g., Enrici, et al., 2011; Lindenberg, et al., 2012; Villarreal, et al., 2008). One study even implicated responses in these areas when a gesture was used to communicate intention (Enrici, et al., 2011).

Yet, a direct investigation into whether there are brain responses that generalize across hand actions with different goals (e.g., as symbolic expressions or to use objects) was previously missing. The current study fills this gap. Here, our main interest was not in characterizing

possible differences. Instead, we examined brain responses for their possible convergence when people view emblems and grasping. Indeed, we identified a substantial amount of overlap. Thus, for manual actions used either to express symbolic meaning or manipulate an object, parietal and premotor responses appear to be non-specific. This places further importance on these regions' more general function in action recognition.

4.3. Supramarginal Divergence

We found that SMG responses in the left hemisphere were strongest to audiovisual speech and weakest to grasping. Conversely, we found that SMG responses in the right hemisphere were strongest to grasping and weakest to speech. In both of these regions, the magnitude of responses to emblems was between that evoked for grasping and speech. The responses to emblems also did not significantly differ from the responses for grasping or speech. These results are consistent with prior findings in tasks other than gesture processing. For example, left SMG is involved when people observe audiovisual speech (Callan, Callan, Kroos, & Vatikiotis-Bateson, 2001; Calvert & Campbell, 2003; Dick, Solodkin, & Small, 2010; Hasson, et al., 2007). In contrast, right SMG is associated with visuo-spatial processing (Chambers, Stokes, & Mattingley, 2004). In particular, numerous studies have implicated this region when people discriminate hand (Ohgami, Matsuo, Uchida, & Nakai, 2004) and finger (Hermsdorfer, et al., 2001) actions, including grasping (Perani, et al., 2001). Thus, our findings in the supramarginal gyri have a different character than those in the frontal and temporal regions, or parietal and premotor regions. In other words, our results suggest that the SMG is sensitive to the sensory quality of the stimuli, rather than to their meaning or to their motor quality.

4.4. Limitations

Because our focus was on the overlap in brain activity when people observe emblems, speech, and grasping, we first determined activity for each against rest. We were then able to characterize convergence between conditions, including even basic commonalities (e.g., in visual cortices; Figure 4).

At the same time, by using a low level resting baseline, we could have also captured non-specific brain activity. In other words, beyond brain activity for processing symbolic and manual features, our functional profiles could include some ancillary activity, not particular to the features of interest.

However, several factors strengthen our ultimate conclusions, despite the fact that we did not use a high level baseline. Most notably, as we discuss above, our results widely corroborate numerous findings. In addition, similar materials (Dick, et al., 2009; Straube, et al., 2012; Xu, et al., 2009) and methods (Dick, et al., 2009; J. I. Skipper, van Wassenhove, Nusbaum, & Small, 2007) have been successfully used to investigate related issues and questions. This includes previous gesture and language experiments that have also used a resting baseline.

Finally, it is worth reiterating that the baseline choice is not trivial. We recognize that careful experimental methodology is especially important for studying the neurobiology of gesture and language (Andric & Small, 2012), given their dynamic relationship in communication and expressing meaning (Kendon, 1994; McNeill, 1992, 2005).

5. SUMMARY

Processing emblematic gestures involves two types of brain responses. One type corresponds to processing meaning in language. The other corresponds to processing hand actions. In this study, we identify lateral temporal and inferior frontal areas that respond

when meaning is presented, regardless of whether that meaning is conveyed by speech or manual action. We also identify parietal and premotor areas that respond when hand actions are presented, regardless of the action's symbolic or object-directed purpose. In addition, we find that the supramarginal gyrus shows sensitivity to the stimuli's sensory modality. Overall, our findings suggest that overlapping, but distinguishable, brain responses coordinate perceptual recognition and interpretation of emblematic gestures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

Andric M, Small SL. Gesture's Neural Language. Front Psychol. 2012; 3:99. [PubMed: 22485103]

Argall BD, Saad ZS, Beauchamp MS. Simplified intersubject averaging on the cortical surface using SUMA. Human Brain Mapping. 2006; 27:14–27. [PubMed: 16035046]

Bartolo A, Cubelli R, Della Sala S, Drei S. Pantomimes are special gestures which rely on working memory. Brain Cogn. 2003; 53:483–494. [PubMed: 14642299]

Bernardis P, Gentilucci M. Speech and gesture share the same communication system. Neuropsychologia. 2006; 44:178–190. [PubMed: 16005477]

Binder JR, Desai RH, Graves WW, Conant LL. Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. Cerebral Cortex. 2009; 19:2767–2796. [PubMed: 19329570]

Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T. Human brain language areas identified by functional magnetic resonance imaging. J Neurosci. 1997; 17:353–362. [PubMed: 8987760]

Binkofski F, Buccino G. Motor functions of the Broca's region. Brain Lang. 2004; 89:362–369. [PubMed: 15068919]

Brennan J, Nir Y, Hasson U, Malach R, Heeger DJ, Pylkkanen L. Syntactic structure building in the anterior temporal lobe during natural story listening. Brain Lang. 2010

Broca P. Remarques sur le siège de la facult e du langage articul e; suivies d'une observation d'aphemie. Bulletin de la Société Anatomique. 1861; 6:330–357.

Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, Seitz RJ, Zilles K, Rizzolatti G, Freund HJ. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. European Journal of Neuroscience. 2001; 13:400–404. [PubMed: 11168545]

Buccino G, Vogt S, Ritzl A, Fink GR, Zilles K, Freund HJ, Rizzolatti G. Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. Neuron. 2004; 42:323–334. [PubMed: 15091346]

Callan DE, Callan AM, Kroos C, Vatikiotis-Bateson E. Multimodal contribution to speech perception revealed by independent component analysis: a single-sweep EEG case study. Brain Res Cogn Brain Res. 2001; 10:349–353. [PubMed: 11167060]

Calvert GA, Campbell R. Reading speech from still and moving faces: The neural substrates of visible speech. Journal of Cognitive Neuroscience. 2003; 15:57–70. [PubMed: 12590843]

Chambers CD, Stokes MG, Mattingley JB. Modality-specific control of strategic spatial attention in parietal cortex. Neuron. 2004; 44:925–930. [PubMed: 15603736]

Chao LL, Haxby JV, Martin A. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. Nat Neurosci. 1999; 2:913–919. [PubMed: 10491613]

Cox RW. AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages. Computers and Biomedical Research. 1996; 29:162–173. [PubMed: 8812068]

Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis: I. Segmentation and surface reconstruction. NeuroImage. 1999; 9:179–194. [PubMed: 9931268]

Dapretto M, Bookheimer SY. Form and content: dissociating syntax and semantics in sentence comprehension. Neuron. 1999; 24:427–432. [PubMed: 10571235]

Decety J, Grezes J, Costes N, Perani D, Jeannerod M, Procyk E, Grassi F, Fazio F. Brain activity during observation of actions - Influence of action content and subject's strategy. Brain. 1997; 120:1763–1777. [PubMed: 9365369]

Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage. 2006; 31:968–980. [PubMed: 16530430]

Devlin JT, Matthews PM, Rushworth MFS. Semantic processing in the left inferior prefrontal cortex: A combined functional magnetic resonance imaging and transcranial magnetic stimulation study. Journal of Cognitive Neuroscience. 2003; 15:71–84. [PubMed: 12590844]

di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G. Understanding Motor Events - a Neurophysiological Study. Experimental Brain Research. 1992; 91:176–180.

Dick AS, Goldin-Meadow S, Hasson U, Skipper JI, Small SL. Co-speech gestures influence neural activity in brain regions associated with processing semantic information. Human Brain Mapping. 2009; 30:3509–3526. [PubMed: 19384890]

Dick AS, Solodkin A, Small SL. Neural development of networks for audiovisual speech comprehension. Brain Lang. 2010; 114:101–114. [PubMed: 19781755]

Duvernoy, HM. The human brain: Structure, three-dimensional sectional anatomy and MRI. Springer-Verlag; New York: 1991. (Chapter Chapter)

Ekman P, Friesen WV. The repertoire of nonverbal communication: Categories, origins, usage, and coding. Semiotica. 1969; 1:49–98.

Enrici I, Adenzato M, Cappa S, Bara BG, Tettamanti M. Intention processing in communication: a common brain network for language and gestures. J Cogn Neurosci. 2011; 23:2415–2431. [PubMed: 20954937]

Fabbri-Destro M, Rizzolatti G. Mirror neurons and mirror systems in monkeys and humans. Physiology. 2008; 23:171–179. [PubMed: 18556470]

Fiebach CJ, Friederici AD, Muller K, von Cramon DY. fMRI evidence for dual routes to the mental lexicon in visual word recognition. J Cogn Neurosci. 2002; 14:11–23. [PubMed: 11798383]

Fiez JA, Raichle ME, Balota DA, Tallal P, Petersen SE. PET activation of posterior temporal regions during auditory word presentation and verb generation. Cereb Cortex. 1996; 6:1–10. [PubMed: 8670633]

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron. 2002; 33:341–355. [PubMed: 11832223]

Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis: II: Inflation, flattening, and a surface-based coordinate system. NeuroImage. 1999; 9:195–207. [PubMed: 9931269]

Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM. Automatically parcellating the human cerebral cortex. Cerebral Cortex. 2004; 14:11–22. [PubMed: 14654453]

Fogassi L, Gallese V, Fadiga L, Rizzolatti G. Neurons responding to the sight of goal directed hand/arm actions in the parietal area PF (7b) of the macaque monkey. Society for Neuroscience. 1998; Vol. 24:257.

Friederici AD, Meyer M, von Cramon DY. Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. Brain Lang. 2000; 74:289–300. [PubMed: 10950920]

Friederici AD, Opitz B, Cramon D. Y. v. Segregating semantic and syntactic aspects of processing in the human brain: An fMRI investigation of different word types. Cerebral Cortex. 2000; 10:698–705. [PubMed: 10906316]

Gallese V, Fadiga L, Fogassi L, Rizzolatti G. Action recognition in the premotor cortex. Brain. 1996; 119:593–609. [PubMed: 8800951]

Gentilucci M, Bernardis P, Crisi G, Dalla Volta R. Repetitive transcranial magnetic stimulation of Broca's area affects verbal responses to gesture observation. Journal of Cognitive Neuroscience. 2006; 18:1059–1074. [PubMed: 16839281]

Gold BT, Balota DA, Jones SJ, Powell DK, Smith CD, Andersen AH. Dissociation of automatic and strategic lexical-semantics: functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. J Neurosci. 2006; 26:6523–6532. [PubMed: 16775140]

Gold BT, Balota DA, Kirchhoff BA, Buckner RL. Common and dissociable activation patterns associated with controlled semantic and phonological processing: Evidence from fMRI adaptation. Cerebral Cortex. 2005; 15:1438–1450. [PubMed: 15647526]

Goldin-Meadow S. The role of gesture in communication and thinking. Trends Cogn Sci. 1999; 3:419–429. [PubMed: 10529797]

Goldin-Meadow, S. Hearing gesture : how our hands help us think. Belknap Press of Harvard University Press; Cambridge, Mass.: 2003. (Chapter Chapter)

Green A, Straube B, Weis S, Jansen A, Willmes K, Konrad K, Kircher T. Neural Integration of Iconic and Unrelated Coverbal Gestures: A Functional MRI Study. Human Brain Mapping. 2009; 30:3309–3324. [PubMed: 19350562]

Grezes J, Armony JL, Rowe J, Passingham RE. Activations related to "mirror" and "canonical" neurones in the human brain: an fMRI study. Neuroimage. 2003; 18:928–937. [PubMed: 12725768]

Hasson U, Skipper JI, Nusbaum HC, Small SL. Abstract coding of audiovisual speech: Beyond sensory representation. Neuron. 2007; 56:1116–1126. [PubMed: 18093531]

Hermsdorfer J, Goldenberg G, Wachsmuth C, Conrad B, Ceballos-Baumann AO, Bartenstein P, Schwaiger M, Boecker H. Cortical correlates of gesture processing: clues to the cerebral mechanisms underlying apraxia during the imitation of meaningless gestures. Neuroimage. 2001; 14:149–161. [PubMed: 11525324]

Hickok G, Poeppel D. The cortical organization of speech processing. Nature Reviews Neuroscience. 2007; 8:393–402.

Hubbard AL, Wilson SM, Callan DE, Dapretto M. Giving speech a hand: Gesture modulates activity in auditory cortex during speech perception. Human Brain Mapping. 2009; 30:1028–1037. [PubMed: 18412134]

Humphries C, Binder JR, Medler DA, Liebenthal E. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. Journal of Cognitive Neuroscience. 2006; 18:665–679. [PubMed: 16768368]

Humphries C, Love T, Swinney D, Hickok G. Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. Human Brain Mapping. 2005; 26:128–138. [PubMed: 15895428]

Indefrey P, Levelt WJ. The spatial and temporal signatures of word production components. Cognition. 2004; 92:101–144. [PubMed: 15037128]

Kalenine S, Buxbaum LJ, Coslett HB. Critical brain regions for action recognition: lesion symptom mapping in left hemisphere stroke. Brain. 2010; 133:3269–3280. [PubMed: 20805101]

Kelly SD, Barr DJ, Church RB, Lynch K. Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. Journal of Memory and Language. 1999; 40:577–592.

Kendon A. Do gestures communicate? A review. Research on Language and Social Interaction. 1994; 27:175–200.

Kircher T, Straube B, Leube D, Weis S, Sachs O, Willmes K, Konrad K, Green A. Neural interaction of speech and gesture: Differential activations of metaphoric co-verbal gestures. Neuropsychologia. 2009; 47:169–179. [PubMed: 18771673]

Knutson KM, McClellan EM, Grafman J. Observing social gestures: an fMRI study. Exp Brain Res. 2008; 188:187–198. [PubMed: 18483724]

Leiguarda RC, Marsden CD. Limb apraxias: higher-order disorders of sensorimotor integration. Brain. 2000; 123(Pt 5):860–879. [PubMed: 10775533]

Lindenberg R, Uhlig M, Scherfeld D, Schlaug G, Seitz RJ. Communication with emblematic gestures: shared and distinct neural correlates of expression and reception. Hum Brain Mapp. 2012; 33:812–823. [PubMed: 21484956]

Lotze M, Heymans U, Birbaumer N, Veit R, Erb M, Flor H, Halsband U. Differential cerebral activation during observation of expressive gestures and motor acts. Neuropsychologia. 2006; 44:1787–1795. [PubMed: 16730755]

Lui F, Buccino G, Duzzi D, Benuzzi F, Crisi G, Baraldi P, Nichelli P, Porro CA, Rizzolatti G. Neural substrates for observing and imagining non-object-directed actions. Soc Neurosci. 2008; 3:261–275. [PubMed: 18979380]

Manthey S, Schubotz RI, von Cramon DY. Premotor cortex in observing erroneous action: an fMRI study. Brain Res Cogn Brain Res. 2003; 15:296–307. [PubMed: 12527103]

Martin A, Chao LL. Semantic memory and the brain: structure and processes. Curr Opin Neurobiol. 2001; 11:194–201. [PubMed: 11301239]

Mason MF, Banfield JF, Macrae CN. Thinking about actions: the neural substrates of person knowledge. Cereb Cortex. 2004; 14:209–214. [PubMed: 14704218]

Mason MF, Norton MI, Van Horn JD, Wegner DM, Grafton ST, Macrae CN. Wandering minds: the default network and stimulus-independent thought. Science. 2007; 315:393–395. [PubMed: 17234951]

McNeill, D. Hand and mind: What gestures reveal about thought. University of Chicago Press; Chicago: 1992. (Chapter Chapter)

McNeill, D. Gesture and thought. University of Chicago Press; Chicago: 2005. (Chapter Chapter)

Mechelli A, Gorno-Tempini ML, Price CJ. Neuroimaging studies of word and pseudoword reading: consistencies, inconsistencies, and limitations. J Cogn Neurosci. 2003; 15:260–271. [PubMed: 12676063]

Mesulam, MM. Patterns in behavioral neuroanatomy: association areas, the limbic system, and hemispheric specialization. In: Mesulam, MM., editor. Principles of behavioral neurology. F.A. Davis; Philadelphia: 1985. p. 1-70.

Miller LM, D'Esposito M. Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. Journal of Neuroscience. 2005; 25:5884–5893. [PubMed: 15976077]

Molnar-Szakacs I, Iacoboni M, Koski L, Mazziotta JC. Functional segregation within pars opercularis of the inferior frontal gyrus: Evidence from fMRI studies of imitation and action observation. Cerebral Cortex. 2005; 15:986–994. [PubMed: 15513929]

Montgomery KJ, Isenberg N, Haxby JV. Communicative hand gestures and object-directed hand movements activated the mirror neuron system. Social Cognitive and Affective Neuroscience. 2007; 2:114–122. [PubMed: 18985130]

Nichols T, Brett M, Andersson J, Wager T, Poline J-B. Valid conjunction inference with the minimum statistic. NeuroImage. 2005; 25:653–660. [PubMed: 15808966]

Nichols T, Holmes AP. Nonparametric permutation tests for functional neuroimaging: A primer with examples. Human Brain Mapping. 2002; 15:1–25. [PubMed: 11747097]

Noll DC, Cohen JD, Meyer CH, Schneider W. Spiral K-space MR imaging of cortical activation. Journal of Magnetic Resonance Imaging. 1995; 5:49–56. [PubMed: 7696809]

Noppeney U, Price CJ. Retrieval of abstract semantics. Neuroimage. 2004; 22:164–170. [PubMed: 15110006]

Ohgami Y, Matsuo K, Uchida N, Nakai T. An fMRI study of tool-use gestures: body part as object and pantomime. Neuroreport. 2004; 15:1903–1906. [PubMed: 15305134]

Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia. 1971; 9:97–113. [PubMed: 5146491]

Patterson K, Nestor PJ, Rogers TT. Where do you know what you know? The representation of semantic knowledge in the human brain. Nat Rev Neurosci. 2007; 8:976–987. [PubMed: 18026167]

Pazzaglia M, Pizzamiglio L, Pes E, Aglioti SM. The sound of actions in apraxia. Curr Biol. 2008; 18:1766–1772. [PubMed: 19013068]

Pazzaglia M, Smania N, Corato E, Aglioti SM. Neural underpinnings of gesture discrimination in patients with limb apraxia. J Neurosci. 2008; 28:3030–3041. [PubMed: 18354006]

Perani D, Fazio F, Borghese NA, Tettamanti M, Ferrari S, Decety J, Gilardi MC. Different brain correlates for watching real and virtual hand actions. Neuroimage. 2001; 14:749–758. [PubMed: 11506547]

Rizzolatti G, Arbib MA. Language within our grasp. Trends in Neurosciences. 1998; 21:188–194. [PubMed: 9610880]

Rizzolatti G, Camarda R, Fogassi L, Gentilucci M, Luppino G, Matelli M. Functional-Organization of Inferior Area-6 in the Macaque Monkey .2. Area F5 and the Control of Distal Movements. Experimental Brain Research. 1988; 71:491–507.

Rizzolatti G, Craighero L. The mirror-neuron system. Annual Review of Neuroscience. 2004; 27:169–192.

Rizzolatti G, Fadiga L, Gallese V, Fogassi L. Premotor cortex and the recognition of motor actions. Cognitive Brain Research. 1996; 3:131–141. [PubMed: 8713554]

Rizzolatti G, Fogassi L, Gallese V. Neurophysiological mechanisms underlying the understanding and imitation of action. Nature Reviews Neuroscience. 2001; 2:661–670.

Rogalsky C, Hickok G. Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. Cereb Cortex. 2009; 19:786–796. [PubMed: 18669589]

Saad, ZS. SUMA: An interface for surface-based intra- and inter-subject analysis with AFNI. IEEE International Symposium on Biomedical Imaging; Arlington, VA. 2004. p. 1510-1513.

Shmuelof L, Zohary E. Dissociation between ventral and dorsal fMRI activation during object and action recognition. Neuron. 2005; 47:457–470. [PubMed: 16055068]

Shmuelof L, Zohary E. A mirror representation of others' actions in the human anterior parietal cortex. J Neurosci. 2006; 26:9736–9742. [PubMed: 16988044]

Skipper JI, Goldin-Meadow S, Nusbaum HC, Small SL. Speech-associated gestures, Broca's area, and the human mirror system. Brain and Language. 2007; 101:260–277. [PubMed: 17533001]

Skipper JI, Goldin-Meadow S, Nusbaum HC, Small SL. Gestures orchestrate brain networks for language understanding. Current Biology. 2009; 19:1–7. [PubMed: 19135370]

Skipper JI, van Wassenhove V, Nusbaum HC, Small SL. Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. Cerebral Cortex. 2007; 17:2387–2399. [PubMed: 17218482]

Stowe LA, Haverkort M, Zwarts F. Rethinking the neurological basis of language. Lingua. 2005; 115:997–1042.

Straube B, Green A, Bromberger B, Kircher T. The differentiation of iconic and metaphoric gestures: Common and unique integration processes. Hum Brain Mapp. 2011; 32:520–533. [PubMed: 21391245]

Straube B, Green A, Jansen A, Chatterjee A, Kircher T. Social cues, mentalizing and the neural processing of speech accompanied by gestures. Neuropsychologia. 2010; 48:382–393. [PubMed: 19782696]

Straube B, Green A, Weis S, Kircher T. A supramodal neural network for speech and gesture semantics: an fMRI study. PLoS One. 2012; 7:e51207. [PubMed: 23226488]

Villarreal M, Fridman EA, Amengual A, Falasco G, Gerscovich ER, Ulloa ER, Leiguarda RC. The neural substrate of gesture recognition. Neuropsychologia. 2008; 46:2371–2382. [PubMed: 18433807]

von Bonin, G.; Bailey, P. The Neocortex of Macaca Mulatta. University of Illinois Press; Urbana: 1947. (Chapter Chapter)

Wagner AD, Pare-Blagoev EJ, Clark J, Poldrack RA. Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. Neuron. 2001; 31:329–338. [PubMed: 11502262]

Wang Y, Ramsey R, de CHAF. The control of mimicry by eye contact is mediated by medial prefrontal cortex. J Neurosci. 2011; 31:12001–12010. [PubMed: 21849560]

Willems RM, Ozyurek A, Hagoort P. When language meets action: The neural integration of gesture and speech. Cerebral Cortex. 2007; 17:2322–2333. [PubMed: 17159232]

Wise RJ, Howard D, Mummery CJ, Fletcher P, Leff A, Buchel C, Scott SK. Noun imageability and the temporal lobes. Neuropsychologia. 2000; 38:985–994. [PubMed: 10775709]

Xu J, Gannon PJ, Emmorey K, Smith JF, Braun A. Symbolic gestures and spoken language are processed by a common neural system. Proceedings of the National Academy of Sciences. 2009; 106:20664–20669.

Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralization of phonetic and pitch discrimination in speech processing. Science. 1992; 256:846–849. [PubMed: 1589767]

HIGHLIGHTS

- We examine brain activity for emblematic gestures, speech, and grasping.

- Gesture and speech activity overlap in lateral temporal and inferior frontal areas.

- Gesture and grasping activity overlap in parietal and premotor areas.

- The left and right supramarginal gyri show opposite speech and grasping effects.

- Brain responses to shared features converge despite different presentation forms.
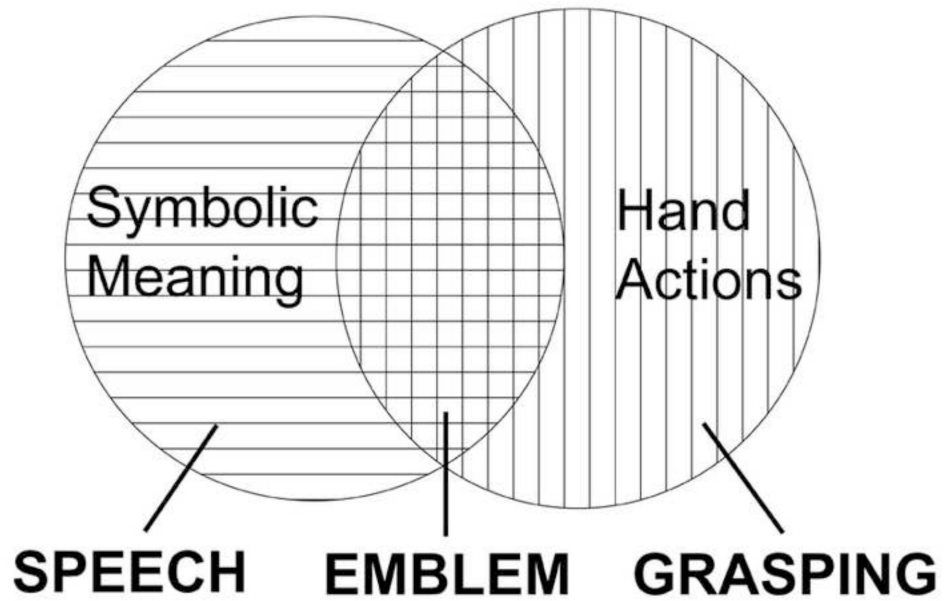
Figure 1.
Conceptual diagram of emblematic gestures (emblems). Emblems share features with speech, since both express symbolic meaning, and with grasping, since both are hand actions.
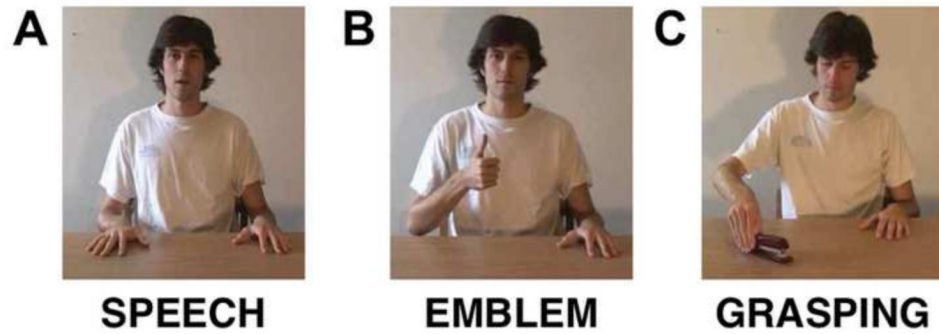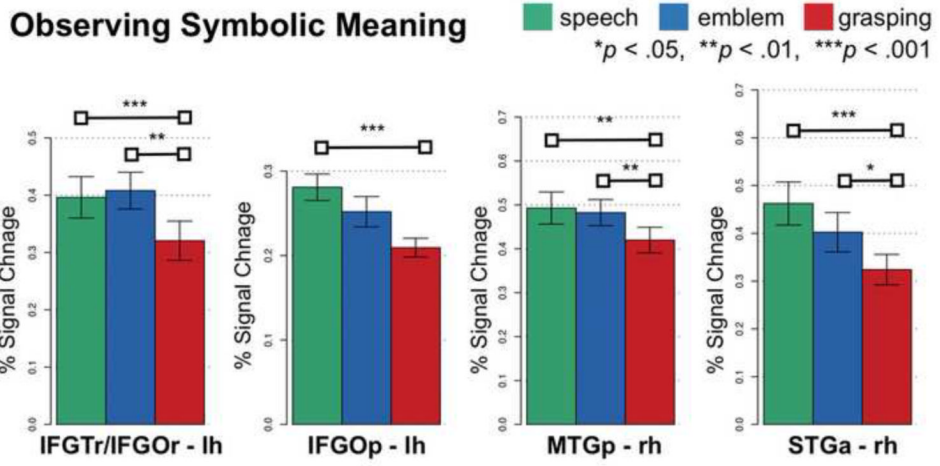
Figure 2.
Still frame examples from videos showing the experimental conditions. (A) Speech, spoken expressions matched in meaning to the emblems. (B) Emblem, symbolic gestures performed with the hand (shown: "it's good"). (C) Grasping, grasping common objects with the hand (e.g., a stapler).
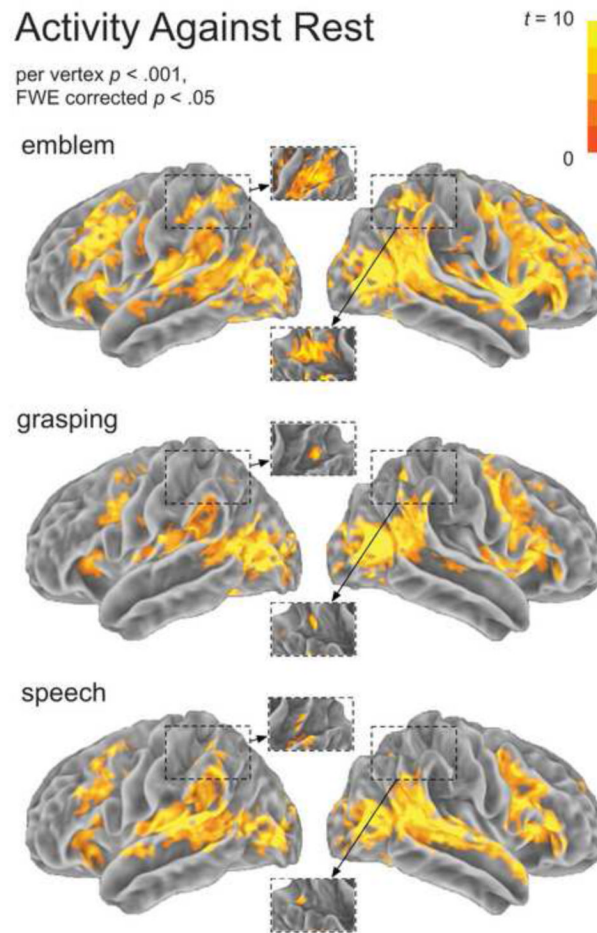
Figure 3.
Activity against rest for each experimental condition. For each condition, the spatial extent of hemodynamic response departures from baseline ("activity") across the cortex is depicted. Insets show the intraparietal sulcus from the superior vantage. The individual per vertex threshold was p < .001, corrected FWE p < .05.
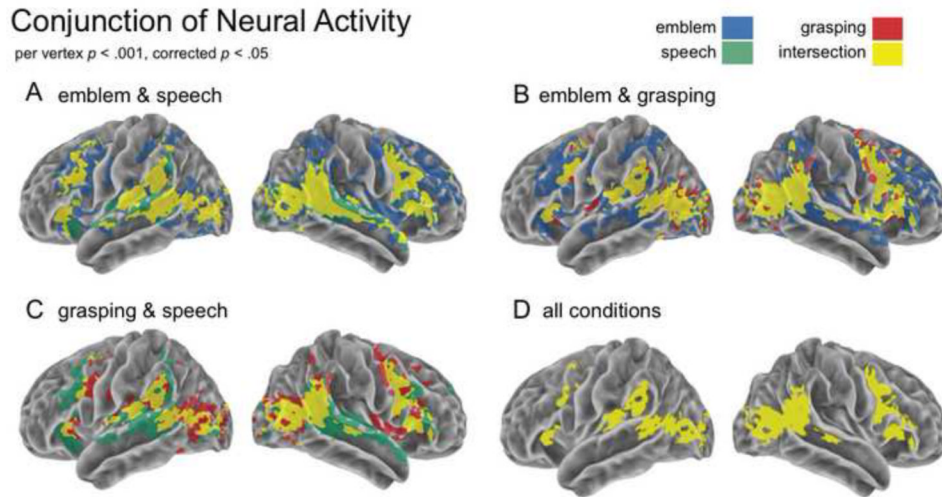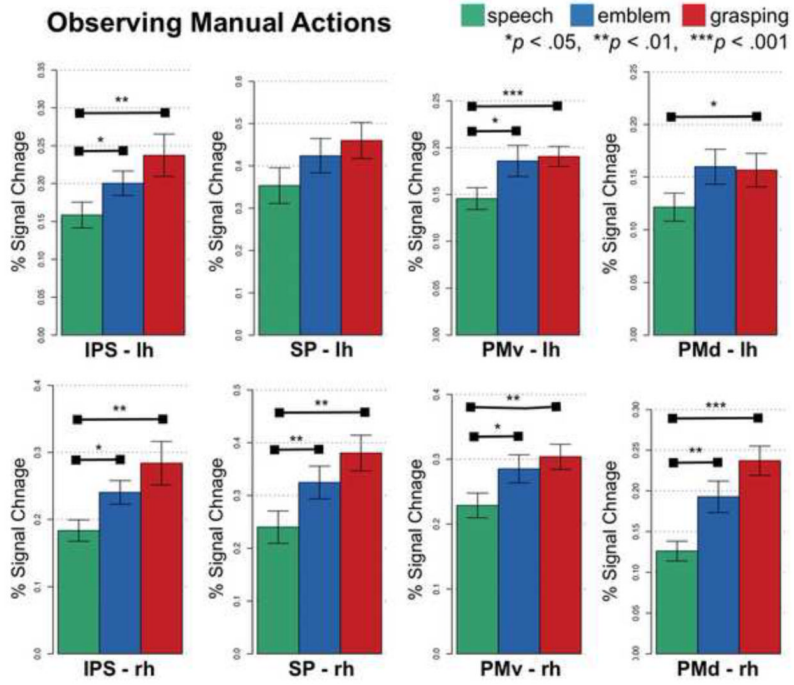
Figure 4.
Conjunction of activity. The overlap of brain activity is shown for each pair of experimental conditions: (A) The spatial extent of overlap highlighting observing symbolic meaning (emblem & speech), (B) purposeful hand actions (emblem & grasping), (C) spoken utterances and grasping (speech & grasping), and (D) all conditions. The individual per vertex threshold was p < .001, corrected FWE p < . 05.
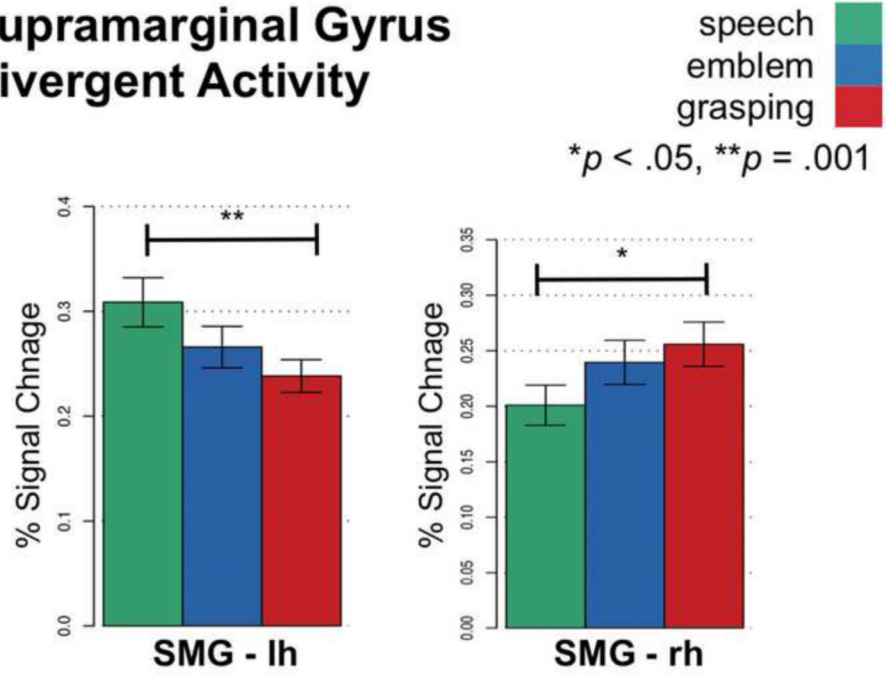
Figure 5.
Regions responsive to observing symbolic meaning. Shown is the neural activity (percent signal change) for each experimental condition in each anatomical region. Horizontal bars connect conditions that significantly differ. Error bars indicate standard error of the mean.

Figure 6.
Regions responsive to observing manual actions. Shown is the neural activity (percent signal change) for each experimental condition in each anatomical region. Horizontal bars connect conditions that significantly differ. Errors bars indicate standard error of the mean.

Figure 7.
Divergent neural activity in left and right supramarginal gyrus. Neural activity in the left SMG was strongest for observing speech and weakest for grasping, whereas in right SMG neural activity was strongest for observing grasping and weakest for speech. Horizontal bars connect conditions that significantly differ. Error bars indicate standard error of the mean.