

Frontal and Temporal Contributions to Understanding the Iconic Co-Speech Gestures That Accompany Speech

Anthony Steven Dick,^{1*} Eva H. Mok,² Anjali Raja Beharelle,^{3,4} Susan Goldin-Meadow,⁴ and Steven L. Small^{1,4,5}

¹Department of Psychology, Florida International University, Florida

²Department of Neurology, The University of Chicago, Illinois

³Department of Psychology, University of Toronto, Ontario, Canada

⁴Department of Psychology, The University of Chicago, Illinois

⁵Department of Neurology, The University of California, Irvine, California



Abstract: In everyday conversation, listeners often rely on a speaker's gestures to clarify any ambiguities in the verbal message. Using fMRI during naturalistic story comprehension, we examined which brain regions in the listener are sensitive to speakers' iconic gestures. We focused on iconic gestures that contribute information not found in the speaker's talk, compared with those that convey information redundant with the speaker's talk. We found that three regions—left inferior frontal gyrus triangular (IFGTr) and opercular (IFGOP) portions, and left posterior middle temporal gyrus (MTGp)—responded more strongly when gestures added information to nonspecific language, compared with when they conveyed the same information in more specific language; in other words, when gesture disambiguated speech as opposed to reinforced it. An increased BOLD response was not found in these regions when the nonspecific language was produced without gesture, suggesting that IFGTr, IFGOP, and MTGp are involved in integrating semantic information across gesture and speech. In addition, we found that activity in the posterior superior temporal sulcus (STSp), previously thought to be involved in gesture-speech integration, was not sensitive to the gesture-speech relation. Together, these findings clarify the neurobiology of gesture-speech integration and contribute to an emerging picture of how listeners glean meaning from gestures that accompany speech. *Hum Brain Mapp* 35:900–917, 2014. © 2012 Wiley Periodicals, Inc.

Key words: gestures; semantic; language; inferior frontal gyrus; posterior superior temporal sulcus; posterior middle temporal gyrus



Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: National Institute of Child Health and Human Development; Contract grant numbers: P01 HD040605, F32 DC008909; Contract grant sponsor: National Institute of Deafness and other Communication Disorders; Contract grant number: RO1 DC003378.

*Correspondence to: Anthony Steven Dick, Department of Psychology, Florida International University, Modesto A. Maidique

Campus, Deuxieme Maison 296B, 11200 S. W. 8th Street, Miami, FL 33199. E-mail: adick@fiu.edu

Received for publication 30 January 2012; Revised 19 September 2012; Accepted 22 October 2012

DOI: 10.1002/hbm.22222

Published online 14 December 2012 in Wiley Online Library (wileyonlinelibrary.com).

INTRODUCTION

Co-speech gestures, the hand and arm movements that speakers routinely produce when they talk, occur naturally in face-to-face communication and play an important role in conveying a speaker's message [Goldin-Meadow, 2003; McNeill, 1992]. The ways in which co-speech gestures contribute to comprehension depend on how they relate to the accompanying speech. For example, beat (rhythmic) gestures have the same form independent of the content of speech and for the most part do not contribute semantic information [McNeill, 1992]. In contrast, iconic gestures, which convey illustrative information about concrete objects or events [McNeill, 1992] and are the focus of our study, function either to support the semantic information available in speech (i.e., they are redundant with speech), or to provide additional semantic information not available in speech [i.e., they are supplementary to speech; see, for example, Alibali and Goldin-Meadow, 1993; Church and Goldin-Meadow, 1986; Perry et al., 1992].

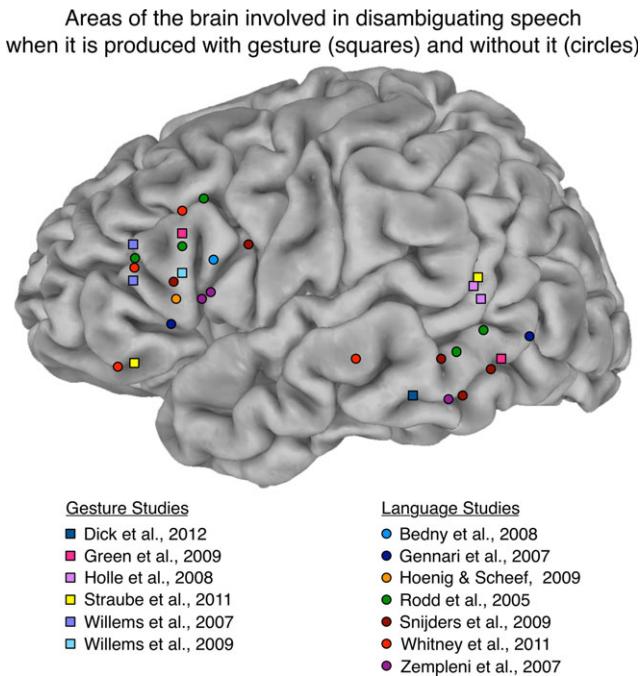
Iconic gestures display in their form an aspect of the object, action, or attribute they represent; e.g., a strumming movement used to describe how a guitar is played [McNeill, 1992]. As a result, iconic gestures can convey information that is also conveyed in the speech they accompany. For example, a speaker might say, "He played the guitar," while making the strumming motion; in this case, gesture is reinforcing the information found in speech. But iconic gestures can also convey information that is not found in the accompanying speech. For example, the speaker might say, "He played the instrument," while producing the same strumming motion; in this case, the speaker's gestures, and only his gestures, tell the listener that the instrument is strummable and thus more likely to be a guitar than a drum, piano, or saxophone. In the second example, fully understanding the speaker's intended message requires the listener to integrate information from separate auditory (speech) and visual (gesture) modalities into a unitary, coherent semantic representation. How the brain accomplishes this integration is the focus of the present study. There is ample behavioral [Beattie and Shovelton, 1999, 2002; Feyereisen et al., 1988; Feyereisen, 2006; Kelly and Church, 1998; Kelly et al., 1999; McNeill et al., 1994] and electrophysiological [Habets et al., 2011; Holle and Gunter, 2007; Kelly et al., 2004; Kelly et al., 2010; Özyürek et al., 2007; Wu and Coulson, 2007a; 2007b] evidence that redundant and supplementary gestures contribute to semantic processing during speech comprehension, and that speakers often use gestures that have the potential to clarify for the listener ambiguities in their speech [Holler and Beattie, 2003]. Electrophysiological evidence also suggests that the way in which gesture relates to speech semantically influences how the brain responds. For example, iconic gestures that support a subordinate meaning of a spoken homonym elicit a stronger event-related potential N400 component (thought to reflect semantic integration processes; Kutas and Federmeier,

2011) in the listener than iconic gestures supporting a dominant meaning. However, the specific brain regions involved in gesture-speech integration are still not well established.

A small number of functional imaging studies have implicated several brain regions that might be involved in gesture-speech integration, including the inferior frontal gyrus (IFG; particularly the anterior *pars triangularis* but also more posterior *pars opercularis*; IFGTr and IFGOP), posterior superior temporal sulcus (STS_p), and posterior middle temporal gyrus (MTG_p; Dick et al., 2009; 2012; Green et al., 2009; Holle et al., 2008; 2010; Kirchner et al., 2009; Skipper et al., 2007; 2009; Straube et al., 2009; 2010; 2011; Willems et al., 2007; 2009; Wilson et al., 2008]. There is, however, little consensus about the nature of the participation of these brain regions in gesture-speech integration.

Of those brain regions implicated in gesture-speech integration, two regions in particular—the IFG and MTG_p—are thought to be critical parts of the distributed network for processing semantic information during language comprehension without gesture [Binder et al., 2009; Lau et al., 2008; Price, 2010; Van Petten and Luka, 2006; Vigneau et al., 2006 for review]. Figure 1 portrays areas of the brain that have been implicated in disambiguating speech when it is produced with gesture (squares) and without it (circles). Across both the language and gesture literatures, increased semantic load appears to result in increased activity in these regions. The figure displays the regions that respond to increased semantic processing demand during comprehension of ambiguous speech (e.g., ambiguous vs. unambiguous words or sentences; subordinate vs. dominant concepts of homonyms; high vs. low demand for semantic retrieval) when it is produced without gesture, and when it is produced with gesture (e.g., gestures that disambiguate the meaning of a homonym; gestures that are semantically incongruent with or unrelated to the sentence context). The figure suggests that two regions in particular—left IFG and left MTG_p (and, in some cases, the right hemisphere homologues of these areas, not shown in Fig. 1)—participate in processing semantic information from both speech and gesture.

With respect to the IFG, Willems et al. [2007] were the first to suggest a role for this region in processing semantic incongruences between iconic gestures and speech. In that study, the authors used stimuli in which the iconic gesture provided information that was either incongruent or congruent with the information conveyed in speech; for example, the verb "wrote" accompanied by a hit gesture (incongruent) or a write gesture (congruent). The authors found that the incongruent conditions elicited greater activity than the congruent conditions in the left IFGTr, a brain region known to be sensitive to increased semantic processing load and conflict [Lau et al., 2008]. Subsequent studies using similar manipulations have also implicated the IFG in gesture-speech integration. In addition to responding more strongly to incongruent iconic gestures

**Figure 1.**

Activation peaks in left inferior frontal and posterior temporal cortex from studies exploring how semantic ambiguity is resolved during language comprehension (e.g., ambiguous vs. unambiguous words or sentences; subordinate vs. dominant concepts of homonyms; high vs. low demand for semantic retrieval) and from studies exploring how gesture contributes to that resolution (e.g., gestures disambiguating the meaning of a homonym; gestures semantically incongruent with or unrelated to the sentence context). Circles indicate language studies and squares indicate gesture studies. Peaks are mapped to a surface representation of the Colin27 brain in Talaraich space.

than to congruent iconic gestures [Willems et al., 2007, 2009], IFG also responds more strongly to metaphoric gestures than to iconic gestures accompanying the same speech [Straube et al., 2011], and to iconic gestures that are unrelated to the accompanying speech than to iconic gestures that are related to the speech [Green et al., 2009]; in other words, to contexts that require additional semantic processing.

However, greater activation in the left IFG has not been found in all studies of gesture-speech integration. For example, we [Dick et al., 2009] found greater responses in the right IFG, but not the left IFG, to hand movements that were not related to the accompanying speech than to iconic and metaphoric gestures that were meaningfully related to the speech. As another example, in a developmental study of 8- to 11-year-old children, we [Dick et al., 2012] found that the activation difference between meaningful and nonmeaningful hand movements was moderated by age in the right, but not the left, IFG. Finally, Holle et al. [2008] found no evidence of involvement of ei-

ther the left or the right IFG in gesture-speech integration; the IFG did not respond differently to gesture that supported the dominant vs. the subordinate meaning of a homonym, nor did it show particular sensitivity to whether the hand movement was an iconic gesture or a grooming movement. Thus, the precise contribution of IFG to gesture-speech integration remains elusive.

Researchers have also examined the role of posterior temporal regions—STSp and MTGp in particular—in the semantic integration of gesture and speech. STSp is more responsive to speech that is accompanied by gesture than to speech alone [Dick et al., 2009], perhaps reflecting this region's putative involvement in processing biologically relevant motion [Beauchamp et al., 2003; Dick et al., 2009; Grossman et al., 2000]. However, Holle et al. [2008, 2010] have suggested that activity in this area reflects sensitivity to the semantic content of gesture, not to biological motion. In their 2008 study, STSp was more active for speech accompanied by meaningful gestures than to speech accompanied by nonmeaningful self-adaptive movements, both of which involve biological motion. Straube et al. [2011] also found increased activation in this region¹ in response to speech accompanied by metaphoric gestures, but not iconic gestures. Finally, Willems et al. [2009] found that the left STSp (and the left MTGp) responded more to speech accompanied by incongruent pantomimes than to the same speech accompanied by congruent pantomimes; they suggest that this region is involved in mapping the information conveyed in gesture and speech onto a common object representation in long-term memory.

Again, greater activation in the STSp has not been found in all studies of gesture-speech integration [e.g., Dick et al., 2009; Green et al., 2009; Kircher et al., 2009; Straube et al., 2009; Willems et al., 2007]. Generally, these studies have found that the MTGp, which is anatomically close to the STSp, is implicated in this function. For example, Green et al. [2009] found that, in German speakers, the left MTGp responded more strongly to German sentences accompanied by unrelated gestures than to the same sentences accompanied by related gestures [although see Willems et al., 2007, 2009].

Present Study

These studies underscore two significant problems facing researchers interested in how the brain accomplishes gesture-speech integration. The first problem is to make sense of the heterogeneous and conflicting set of regions that are implicated in gesture-speech integration at the

¹We note that in their Table III, Straube et al. (2011) label this activation centered at MNI $x = -56, y = -52, z = 12$ in the left middle temporal gyrus. However, independent verification of the MNI coordinates and inspection of their Figure 3 with reference to a published atlas [Duvernoy, 1999; Mai et al., 2007] shows that the activity is clearly in the left superior temporal sulcus (also see our Fig. 1).

**Conditions: 2 (Gesture: Present vs Absent)
x 2 (Language: Specific vs Nonspecific)**

Specific/Nonspecific Language Embedded in Narrative

Stacie was excited when her parents brought home her new **parrot/pet**.
 Sparky, who was cute as could be, tried to bite her when she fed him.
 The mess that he made all over forced Stacie to put down newspapers.
 When she tried to move him onto the newspapers, Sparky **pecked/attacked** her again!
 That night, as Stacie tried to put Sparky in a cage, he **flapped/struggled** non-stop
 and even scratched her with his claws!
 She must have left the cage open, because in the morning Sparky had escaped.



Gesture+Specific Language



No Gesture+Specific Language

Specific Language
That night, as Stacie tried to put Sparky in a cage, he **flapped** non-stop
and even scratched her with his claws!



Gesture+Nonspecific Language



No Gesture+Nonspecific Language

Nonspecific Language
That night, as Stacie tried to put Sparky in a cage, he **struggled** non-stop
and even scratched her with his claws!

Figure 2.

An example of a narrative used in all four conditions (top). Language specificity was varied in three sentences (see the bolded words): pet, attacked, struggled are the nonspecific terms; parrot, pecked, flapped are the specific terms. Half of the narratives were accompanied by gestures, and half were not, creating four conditions (bottom). The same gestures were used with the

nonspecific and specific language. The gesture was redundant with the specific language (Gesture+Specific Language), but added information to the nonspecific language, rendering it more specific (Gesture+Nonspecific Language). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

semantic level. The second is to provide more anatomical precision in characterizing brain responses, with particular attention to the activation clusters that extend across several regions. The challenge here is not only to refine our understanding of how the brain accomplishes gesture-speech integration, but also to address broader issues, such as how the brain accomplishes semantic processing during language comprehension.

To achieve these ends, our specific goal in the present study is to clarify which of the regions implicated in both gesture-speech integration and in semantic comprehension without gesture (i.e., IFG, STSp, and MTGp) are specifically sensitive to the semantic relation between iconic gestures and speech. We address these issues in two ways. First, we use a novel experimental paradigm that probes semantic

integration of gesture and speech during naturalistic language processing, using gestures that add information that has the potential to be integrated with the information in speech. Importantly, the information conveyed in gesture does not contradict the information conveyed in speech. Second, we use surface-based image analysis with anatomically rigorous regional specification. Surface-based analysis improves anatomical registration across participants [Argall et al., 2006; Desai et al., 2005; Fischl et al., 1999; Van Essen et al., 2006], and a anatomical regions of interest (ROI) approach enables us to define the response in anterior and posterior IFG, STSp, and MTGp at the level of the individual participant [Devlin and Poldrack, 2007].

To identify those regions involved in gesture-speech semantic integration, we employed a 2×2 factorial design

(see Fig. 2). In the first factor, Gesture, story narratives were either accompanied by an iconic gesture (Gesture) or they were not (No Gesture). In the second factor, Language Specificity, the lexical meaning of the target word in speech was either specific (Specific Language, e.g., in "he flapped nonstop," the word "flapped" specifies a manner that is characteristic of a particular type of object, in this case, a bird) or not specific (Nonspecific Language, e.g., in "he struggled nonstop," the word "struggled" provides less information about the particular object doing the action than does "flapped"). We crossed these factors and manipulated the informativeness of the iconic gestures by controlling the degree to which the gesture reinforced the meaning in the accompanying speech (e.g., a flapping gesture produced along with the word "flapped," that is, a redundant gesture in the Gesture+Specific Language condition), or provided additional information not found in speech (e.g., a flapping gesture produced along with the word "struggled," that is a supplemental gesture in the Gesture+Nonspecific Language).

We related the participants' processing of the narratives (based on a post-scan recognition test) to the brain response in our defined regions of interest. Based on our review of the literature, we predicted that left IFG and MTGp would respond more strongly to supplemental gestures (i.e., Gesture+Nonspecific Language) than to redundant gestures (i.e., Gesture+Specific Language) because, in each pair, it required more work to integrate the gesture with the nonspecific language than with the specific language. However, when the speech was not accompanied by gesture (i.e., in the No Gesture conditions), there was no gesture information to be integrated with speech, although the language still varied with respect to specificity: "he struggled nonstop" (No Gesture+Nonspecific Language) leaves open who the actor is in a way that "he flapped nonstop" (No Gesture+Specific Language) does not. Thus, if left IFG and MTGp are responsive to the amount of work needed to integrate the information in gesture and speech (and not merely to the level of specificity in the speech), the two regions should respond more strongly in the Nonspecific Language vs. Specific Language conditions only when they are produced with gesture, and not when they are produced without gesture. Thus, we predicted a 2×2 interaction in the left IFG and MTGp regions. Because we have not found STSp to be involved in gesture-speech integration at the semantic level in our previous work [Dick et al., 2009, 2012], we did not expect STSp to respond to the semantic manipulation, and thus predicted no interaction for these regions.

MATERIALS AND METHODS

Participants

Seventeen right-handed native speakers of American English (11 F; M = 22.4 years; SD = 4.64 years; range 18–34 years) participated. Each gave written informed consent

following the guidelines of Institutional Review Board for the Division of Biological Sciences of The University of Chicago, which approved the study. All participants reported normal hearing and normal or corrected-to-normal vision. None reported any history of neurological or developmental disorders. Three additional participants were excluded for failure to complete the study ($n = 2$), and for technical problems with the scanner ($n = 1$). Four additional participants performed worse than chance on the post-test (i.e., fewer than 7 out of 16 correct), which was taken as evidence that they did not pay attention to the stories during the scanning session. Because we do not request overt decisions or motor responses during fMRI studies [Small and Nusbaum, 2004], we rely on such post-tests to ensure compliance with task performance.

Image Acquisition

Imaging data were acquired on a 3T Siemens Trio scanner. A T1-weighted structural scan was acquired before the functional runs for each participant (1 mm \times 1 mm \times 1 mm resolution; sagittal acquisition). Gradient echo echo-planar T2* images optimized for blood oxygenation level dependent (BOLD) effects were acquired in 32 axial slices with an in-plane resolution of 1.7 mm \times 1.7 mm, and a 3.4 mm slice thickness with 1 mm gap (TR/TE = 2,000/20 ms, Flip Angle = 75°). Four dummy volumes at the beginning of each run were acquired and discarded.

Materials

The primary cognitive manipulation was the specificity of the verbal information (Specific Language vs. Nonspecific Language), and whether the sentence was accompanied by an iconic gesture or not (Gesture vs. No Gesture). Thus, in some stories (Gesture + Specific Language), the iconic gestures were redundant with the verbal information and in others (Gesture+Nonspecific Language) the same iconic gestures contributed supporting information to the verbal information. The stories across conditions were identical except for three words within each story. That is, nonspecific words in one version (e.g., pet, attacked, struggled), which allowed for multiple interpretations, were replaced with specific words in the other version (e.g., parrot, pecked, flapped), which allowed only a single unique interpretation (see Fig. 2 and Supporting Information Materials for examples). This design established the first experimental factor, Language Specificity, with two levels (Specific Language vs. Nonspecific Language). The second experimental factor, Gesture, also had two levels. Here, the narratives were either accompanied by meaningful iconic gestures (Gesture), or by no hand movements (No Gesture). We also included a fifth condition in which rhythmic beat gestures were produced along with Specific Language (Beat Gesture+Specific Language). Beat gestures do not convey substantive information but rather

serve a discourse function [McNeill, 1992]. This condition allowed us to identify brain regions that respond to naturalistic gestures that accompany speech, regardless of whether they contribute semantic information. We report the analysis of this condition in Supporting Materials and do not discuss it further in the article.

For each condition, video recordings were made of a male actor (a native speaker of American English) performing each narrative. The actor rehearsed each narrative until he was fluent, and until he could keep gestures and prosody as identical as possible across conditions. In the No Gesture conditions, the speaker held his hands at his side. In the Gesture conditions, he produced three iconic gestures over the course of the entire story conveying specific information in each narrative (e.g., a flapping motion, a flying motion, and then another flapping motion, all of which pinpointed and consistently reinforced a bird as the focus of the narrative). The speaker produced the iconic gestures naturally with speech (i.e., the timing of the gesture with speech was not edited in any way; cf. Holle et al., 2008). Each video was edited to $30\text{ s} \pm 1.5\text{ s}$ in length using Final Cut Pro (Apple Inc, Cupertino, CA) and the sound volume was normalized.

It is important to point out that narratives in the Gesture+Specific Language, Gesture+Nonspecific Language, No Gesture+Specific Language conditions all allowed the same semantic interpretation—that the story was about a bird. This information was conveyed in both gesture and speech in the Gesture+Specific Language condition, only in gesture in the Gesture+Nonspecific Language condition, and only in speech in the No Gesture+Specific Language condition. Note that without the iconic gesture to specify a particular target, the language in the No Gesture+Nonspecific Language condition left open a variety of interpretations, only one of which was consistent with the interpretation of the story in the other three conditions (e.g., a participant listening to the No Gesture+Nonspecific Language narrative displayed in Figure 2 might guess that the story was about a puppy rather than a parrot).

A total of 20 narratives were constructed, and each was recorded in all five forms (including the Beat Gesture + Specific Language condition), for a total of 100 distinct video clips. Each narrative contained an average of 88.85 words (range = 74–99 words). The constructed narratives were matched for total word length, syntactic complexity, and average printed word frequency. To determine word frequency, we used the average of the 1st through 6th grade printed word frequency list published in the Educator's Word Frequency Guide (Zeno et al., 1995).

Experimental Procedure

The functional MRI paradigm was a block design, with each narrative comprising one “block” separated by a fixed rest interval of 18 s to allow the hemodynamic response to return to baseline. Four narratives per condi-

tion were presented to each participant, split across two runs (20 total stories, or blocks, per participant). Thus, for the iconic gesture conditions, participants viewed 12 gestures (3 gestures per narrative, across four narratives). Four pseudo-random stimulus sequences were generated such that two narratives per condition were presented in each run and no narratives were repeated within subject. The stimuli were presented using Presentation software (Neurobehavioral Systems, Albany, CA) and projected onto a back-projection screen that participants viewed using a mirror attached to the head coil. Sound was conveyed through MRI compatible headphones, and the sound level was individually adjusted prior to the functional runs to each participant's comfort level.

Participants were instructed to pay careful attention to the narrator and were told that they would be asked questions about the narratives after they came out of the scanner. After the scanning session (15–20 min after the presentation of the second run), participants were given the post-scan recognition test.

Postscan Recognition Test

The post-scan recognition test consisted of a 4-alternative forced choice question about the topic of each narrative. To assess whether participants were paying attention during story presentation, we used the binomial distribution to determine that participants needed to answer 7 or more questions correctly to achieve greater than chance performance ($n = 16$ across the Specific language conditions, including the Beat+Specific Language condition, and the Gesture+Nonspecific Language condition; per-trial probability = 0.25; one-tailed). In this assessment of whether participants were paying attention, we did not include the No Gesture+Nonspecific Language condition, because there was technically no “correct” answer. That is, the language was nonspecific and there was no gesture to specify a particular referent. However, the No Gesture+Nonspecific Language condition does provide a guessing baseline against which the other conditions of interest can be assessed. Thus, for all conditions we tallied the proportion of answers that matched the specific language answer (e.g., referring to Figure 2, participants sensitive to the specific meaning would have answered that the particular pet in the story was a “parrot” instead of “cat,” “dog,” or “hamster”).

For this analysis, the mean proportions for each condition were: Gesture+Specific Language $M = 0.84$; No-Gesture+Specific Language $M = 0.88$; Gesture+Nonspecific Language $M = 0.72$; No Gesture+Nonspecific Language $M = 0.18$, and there were significant differences across the four conditions, Friedman $\chi^2(3) = 30.59$, $P < 0.001$. Post-hoc Wilcoxon matched-pairs signed-ranks tests (FDR corrected) showed differences between the No Gesture+Nonspecific Language condition (i.e., the baseline condition) and the other three conditions: compared to

Gesture+Specific Language, $W = 134.5$, $P < 0.001$, 95% CI 0.50 – 0.88; compared with No Gesture+Specific Language, $W = 136.0$, $P < 0.001$, 95% CI 0.63–0.88; compared with Gesture+Nonspecific, $W = 149.0$, $P < 0.001$, 95% CI 0.38–0.75. There were no reliable differences between pairings of any of the other three conditions. The important point is that, for nonspecific speech, participants were unable to identify a particular referent (i.e., in the No Gesture+Nonspecific Language condition). But when that nonspecific speech was accompanied by gesture (i.e., in the Gesture+Nonspecific Language condition), participants were able to identify the referent that would have been encoded by specific language. These findings thus make it clear that the participants were gleaning information from the gestures they saw and integrating that information with what they heard in speech.

Data Analysis

Post-processing

Post-processing steps conducted with AFNI/SUMA [Cox, 1996; <http://afni.nimh.nih.gov>] in the native volume domain included time series despiking, slice-timing correction, spatial registration of the functional volume to the structural volume, and three-dimensional affine motion correction using weighted least-squares alignment of three translational and three rotational parameters. Individual cortical surfaces for each subject were constructed from the T1 volume using Freesurfer [Dale et al., 1999; Fischl et al., 1999]. The post-processed time series were then projected from the volume domain to the surface domain using AFNI/SUMA and the time series was mean-normalized. Because it can lead to contamination of the time series across regions of interest [ROIs; Neito-Castanon et al., 2003], the time series were not spatially smoothed for the ROI analysis. However, for the whole-brain analysis, spatial smoothing [6 mm FWHM using a HEAT kernel; Chung et al., 2005] of the time series was applied in the surface domain. Note that all other analyses were also performed in the surface domain.

To determine the degree of BOLD activity against a resting baseline, we modeled each narrative as a block. The stimulus presentation design matrix was convolved with a gamma function model of the hemodynamic response, and this served as the primary regressor of interest for each of the five predictors (four conditions plus localizer) in a general linear model. Additional regressors included six estimated motion parameters, the time series mean, and linear and quadratic drift trends. Beta and t -statistics were estimated independently for each surface vertex for each individual participant.

Analysis

The analysis consisted of three parts. At the level of the whole brain, we first established which regions were more

active for iconic gestures compared to no gestures, and which regions were sensitive to the 2×2 Gesture by Language Specificity interaction. We next focused on particular cortical ROIs identified on the basis of the literature, and examined how the BOLD response differed in amplitude across conditions. Here we also focused on the 2×2 Gesture by Language Specificity interaction. Finally, we examined how the signal in these identified ROIs related to post-scan recognition performance.

Whole-Brain Analysis

We examined two focused comparisons at the level of the whole brain: (1) a comparison of the iconic gesture conditions compared with the conditions without gestures (i.e., the main effect of Gesture); (2) a 2 (Language Specificity) \times 2 (Gesture) interaction contrast, to identify regions that respond more strongly to Nonspecific Language than Specific Language only when accompanied by gesture, and not when they occur without gesture. Thus, we conducted a 2×2 repeated measures ANOVA, with both factors defined as repeated measures, for each vertex on the whole brain.

For this analysis, each individual surface was standardized to contain the same number of vertices using icosahedral tessellation and projection [Argall et al., 2006]. Beta weights were mapped from the individual surfaces to the standardized mesh surface where the statistical analysis was carried out. An average group surface was created from the 17 individual cortical surfaces on which to present the group data. All group maps were thresholded at a single vertex threshold of $P < 0.05$ with family-wise error (FWE) correction at the whole-brain level at $P < 0.05$. This FWE correction was determined by defining a minimum cluster area following a Monte Carlo simulation conducted as part of the Freesurfer software package (cf. Forman et al., 1995).

Region of Interest Analysis

We were primarily interested in identifying regions that were sensitive not only to the presence of gesture, but also to the semantic relation between gesture and speech. On the basis of prior literature, we expected the IFG, STSp, and MTG to be the regions most likely to be involved in gesture-speech integration at the semantic level. We defined these regions anatomically on individual cortical surfaces based on manual refinement of the automatic Freesurfer parcellation [Desikan et al., 2006; Fischl et al., 2004], which is itself based on the anatomical conventions of Duvernoy [1999]. Thus, we divided (a) the IFG into an anterior IFGTr and posterior IFGOP, which have a different cytoarchitecture (Amunts et al., 1999) and connectivity profile with the temporal lobe [Saur et al., 2008; Schmahmann and Pandya, 2006]; and (b) the STSp into an upper bank and lower bank, with a boundary at the fundus of

the sulcus. The upper bank of the STSp receives strong projections from both auditory and visual association cortex, while the lower bank is more strongly connected to the middle and inferior temporal gyri and visual cortices [Seltzer and Pandya, 1978, 1989, 1994]; (c) the middle temporal gyrus into anterior and posterior (MTGp) portions. STSp and MTGp were obtained by manually subdividing the standard Freesurfer parcellation into anterior and posterior parts. A vertical plane extending from the anterior tip of the transverse temporal gyrus served as the anterior-posterior dividing line. In summary, the regions examined were (both left and right) IFGOp and IFGTr, the upper bank (STSp_upper) and lower bank (STSp_lower) of the STSp, and the MTGp.

We used the R statistical package (v. 2.13.1; <http://www.R-project.org/>; R Development Core Team, 2011) for the anatomical ROI analysis. For this analysis, hemodynamic response estimates (betas) for each experimental condition were extracted from each anatomical region from each individual participant. Outlying vertices more than three standard deviations from the region mean were discarded. To identify regions that were sensitive not only to the presence of gesture, but also to the semantic relation between gesture and speech, we assessed the 2×2 interaction of Language Specificity (Specific Language vs. Nonspecific Language) by Gesture (Gesture vs. No Gesture) using a linear mixed-effects model. This model allowed the intercepts and slopes to vary across the random effect "subject". To obtain the beta estimates for the interaction, we used restricted maximum likelihood (REML), which takes into account the degrees of freedom of the fixed effects when estimating the variance components. The null hypothesis of the interaction specifies that the addition of iconic gestures should have no effect on the interpretation of the narrative, and thus no effect on the response in brain regions involved in gesture-speech integration (i.e., [Gesture+Nonspecific Language – Gesture+Specific Language] – [No Gesture+Nonspecific Language – No Gesture+Specific Language] = 0, or in other words, the difference of the differences between the conditions should be zero). If there is a significant interaction driven by a greater response to Gesture+Nonspecific Language compared with Gesture+Specific Language, this pattern would indicate that the region responds more strongly to nonspecific than to specific language, but only when that language is accompanied by gesture, which can resolve the ambiguity inherent in the nonspecific language. In other words, as in the whole brain analysis, a significant interaction would suggest that the region is involved in integrating semantic information from gesture and speech.

As recommended by Krishnamoorthy et al. [2007], we calculated standard errors and 95% confidence intervals (CIs) of the beta estimates using a bootstrap approach. To perform the bootstrap, we adapted the procedure from Venables and Ripley [2010, p. 164] to the mixed-effects regression model and resampled the residuals. In this pro-

cedure, the linear mixed effects model is fit to the data, the residuals are resampled with replacement, and new model coefficients are estimated. This process is iterated 5,000 times to define the standard errors of each parameter estimate. The bootstrap standard errors are used to calculate the 95% CIs and significance tests (*t* values).

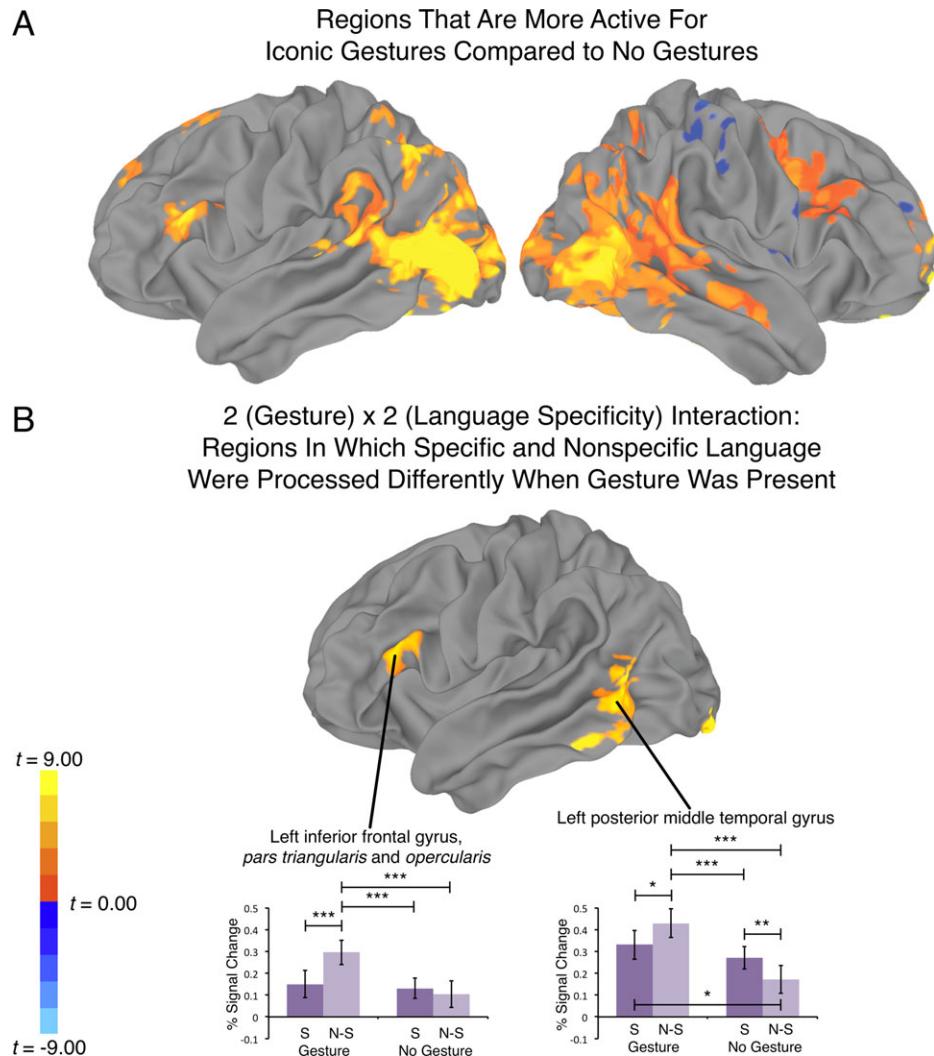
Brain-Behavior Relations With the Post-Scan Recognition Test

To relate brain activity to subsequent recognition memory, we examined the relation between the average hemodynamic response in each of the ten previously defined ROIs (left and right IFGOp, IFGTr, STSp_upper, STSp_lower, MTGp) and participants' post-test scores. We focused on the relation between the signal in these regions and performance on questions related to the Gesture+Nonspecific Language, Gesture+Specific Language, and No Gesture+Specific Language conditions (i.e., No Gesture+Nonspecific Language was not examined because there was no single correct response to the narratives in this condition; see Method). To examine the statistical relationships between regional signal and recognition, in order to model the ordinal response variable, we conducted proportional odds logistic regression using a logit link function [i.e., ordered logit regression; Hardin and Hilbe, 2007]. A FDR procedure [Benjamini and Hochberg, 1995; Genovese et al., 2002] was used to correct for multiple comparisons across ROIs.

RESULTS

Whole Brain Analysis

Figure 3A and Table I show regions more active when iconic gestures accompanied speech compared with when they did not (i.e., the main effect of Gesture). Notably regions of the dorsal "where/how" visual stream are reliably more active for gestures. These regions include bilateral striate and extrastriate visual areas, including the anterior occipital sulcus, putative site of the motion sensitive area V5/MT+ [Malikovic et al., 2007], and the intraparietal sulcus and superior parietal lobe. Greater activity for Gesture was also revealed in the posterior superior temporal sulcus, a region that is responsive to biologically relevant motion [Beauchamp et al., 2003; Grossman et al., 2000; Saygin et al., 2004; Thompson et al., 2005], and the planum temporale and inferior supramarginal gyrus, regions potentially involved in sensorimotor integration [Hickok and Poeppel, 2007]. Gestures also elicited greater activity in the left IFG, and right middle frontal gyrus and inferior frontal and precentral sulcus, while the No Gesture conditions were more active in the right postcentral gyrus, insula, and superior frontal gyrus. Broadly, these findings are consistent with those we reported in a

**Figure 3.**

A: Regions that were more active for conditions with iconic gestures compared to no gestures ($P < 0.05$, corrected). **B:** The 2 (Gesture) \times 2 (Language Specificity) Interaction: Regions in which specific and nonspecific language were processed differently when gesture was present ($P < 0.05$, corrected). For B,

there were no right hemisphere clusters that survived the multiple comparison correction. S, specific language; N-S, nonspecific language; *** $P < 0.001$ (corrected); ** $P < 0.01$ (corrected). * $P < 0.05$ (corrected).

previous study of gestures on a different sample of participants [Dick et al., 2009].

Figure 3B and Table I present regions that show a significant Gesture (2) \times Language Specificity (2) interaction. Two of these regions, IFG and MTGp, were predicted on the basis of our review presented in the Introduction. We interrogated the signal in these regions further by averaging the betas for each condition in surface vertices that showed a significant interaction (Fig. 3B graphs). Both regions showed a significant interaction (IFG; $t(14) = 5.04$, $P < 0.001$, $b = 0.17$; 95% CI = 0.10–0.24; MTGp; $t(14) = 8.26$, $P < 0.001$, $b = 0.20$; 95% CI = 0.15–0.25. Post hoc comparisons across conditions [$P < 0.05$ corrected using

the FDR procedure; Genovese et al., 2002; Benjamini and Hochberg, 1995] are reported in the graphs. Both regions respond more strongly to the Nonspecific Language condition, but only when that language was produced with iconic gesture. These results are consistent with the ROI analysis we report below, and we turn now to discussion of those findings.

ROI Analysis

In the ROI analyses, we aimed to identify regions that are sensitive to the interaction between speech and iconic

TABLE I. Regions showing reliable differences for the main effect of Gesture, and for the Gesture by Language Specificity interaction

Region	Talairach				CS (Area)	MI
	x	y	z	BA		
Iconic Gestures > No Gestures						
L. Middle temporal gyrus/anterior occipital sulcus	-42	-64	7	37	14136 (5,360.21)	0.346
L. Precuneus	-7	-72	35	7	2,054 (771.69)	0.168
L. Inferior frontal gyrus	-54	19	20	45	1,024 (499.94)	0.160
L. Middle frontal gyrus/superior frontal sulcus	-27	19	-42	8	899 (393.55)	0.122
L. Superior frontal gyrus	-13	51	34	9	757 (452.70)	0.124
R. Fusiform gyrus	56	-51	-20	37	17,766 (622.89)	0.514
R. Precentral gyrus/sulcus	49	-1	38	6	2,872 (1018.67)	0.156
R. Posterior cingulate gyrus	8	-45	10	29	1,871 (602.05)	0.195
R. Orbital gyrus/gyrus rectus	12	45	-22	11	1,588 (893.04)	0.998
No gestures > Iconic gestures						
R. Postcentral gyrus	55	-34	49	40	1,430 (439.11)	0.139
R. Insula	48	3	5		846 (292.18)	0.128
R. Superior frontal gyrus	23	51	23	10	506 (307.11)	0.07
Interaction [Gesture: Nonspecific Language–Specific Language] – [No Gesture: Nonspecific Language–Specific Language]						
L. Middle temporal gyrus	-54	-59	1	37	1,276 (520.72)	0.358
L. Lingual gyrus	-14	-95	-12	17	1,039 (508.73)	0.547
L. Inferior frontal gyrus	-53	20	18	45	675 (343.33)	0.230

Individual voxel threshold $p < .05$, corrected (FWE $p < 0.05$). Center of mass defined by Talairach and Tournoux coordinates in the volume space. BA = Brodmann Area. CS = Cluster size in number of surface vertices. Area = Area of cluster. MI = Maximum intensity (in terms of percent signal change).

gestures. We examined 10 regions—left and right IFGOp, IFGTr, MTGp, STSp_upper, STSp_lower—for the presence of a 2×2 Language Specificity (Nonspecific vs. Specific) by Gesture (Gesture vs. No Gesture) interaction. A significant interaction was found in three regions: left IFGTr, left IFGOp, and left MTGp (left IFGTr: $t(14) = 2.19$, $P = 0.045$, $b = 0.08$; 95% CI = 0.01–0.15; left IFGOp: $t(14) = 2.27$, $P = 0.04$; $b = 0.06$; 95% CI = 0.01–0.11; left MTGp: $t(14) = 6.10$, $P < .001$, $b = 0.14$, 95% CI = 0.10–0.19). Inspection of Figure 4 indicates that left IFGTr, IFGOp, and MTGp regions respond more strongly during conditions containing Non-specific Language than during conditions containing Specific Language, but only when that language was produced with iconic gesture and not when it was produced without gesture. Notably, right IFGTr and IFGOp, and left STSp, which have been associated with gesture-speech integration in prior studies [Dick et al., 2009; Holle et al., 2008; Straube et al., 2011], did not show sensitivity to the semantic manipulation in this study.

As an exploratory analysis, we examined all pairwise comparisons across the 10 ROIs (60 total comparisons), and report both corrected ($P < 0.05$ FDR corrected) and uncorrected ($P < 0.05$) comparisons in Figures 4 and 5. We report uncorrected values because multiple comparison procedures guard against committing any Type I error (regardless of whether the null hypothesis is or is not true), and they can be quite restrictive when there are a large number of comparisons [Benjamini and Gavrilov,

2009; Benjamini and Hochberg, 2000]. It is informative to report uncorrected values, but we note that we restrict our interpretation in the Discussion to only those comparisons that were statistically significant after correction. In general, the results of the ROI analysis showed that left IFGTr, IFGOp, and MTGp are sensitive to the semantic information conveyed in iconic gestures, and that these regions are involved in integrating iconic gestures with speech at the semantic level when the speech is ambiguous.

Brain-Behavior Relations With the Post-Scan Recognition Test

To model the ordinal outcome variable of recall, we conducted 30 proportional odds logistic regressions (across the 10 ROIs: left and right IFGOp, IFGTr, MTGp, STSp_upper, STSp_lower). Regional signal during the Gesture+Nonspecific Language, Gesture+Specific Language, and No Gesture+Specific Language conditions comprised the predictors, and percent correct in the post-scan recognition tests for each of these conditions comprised the outcome. In only two regions—left and right STSp, upper bank—was the regional signal a significant predictor of post-scan recognition for stories with gestures; no significant relationships were found for stories without gestures (Fig. 6). For the following comparisons, the 95% confidence intervals of the parameters did not overlap with zero, and the

Left Hemisphere Regions of Interest

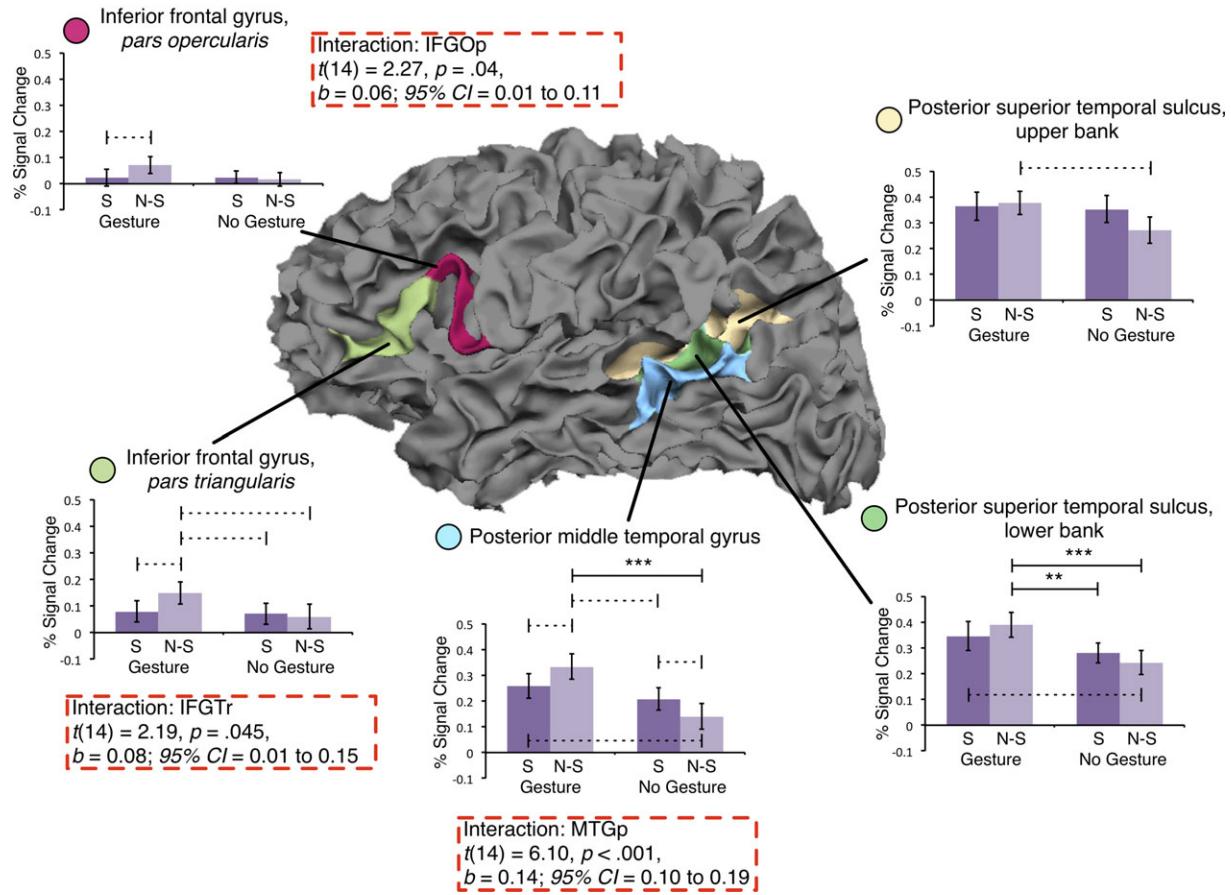


Figure 4.

Results of the analysis for left hemisphere inferior frontal and posterior temporal regions of interest. Regions are outlined on a white matter surface image of a representative subject with the pial surface removed to reveal the sulci. Three regions—IFGOp, IFGTr, and MTGp—showed a significant interaction across the Language Specificity and Gesture factors (dashed red

boxes). We see more activity in response to NonSpecific Language than to Specific Language, but only when the language is accompanied by Gesture (i.e., in the Gesture bars and not in the No Gesture bars). S, specific language; N-S, nonspecific language; *** $P < 0.001$ (corrected); ** $P < 0.01$ (corrected); -- $P < 0.05$ (uncorrected).

parameter estimates were reliable. We report McFadden Pseudo- R^2 values: Gesture+Nonspecific Language, right STSp_upper: $b = 5.22, 95\% \text{ CI} = 1.09\text{--}10.32, t(15) = 2.26, P = 0.04; R^2 = 0.14$; Gesture+Specific Language, left STSp_upper: $b = 11.58, 95\% \text{ CI} = 2.27\text{--}20.88, t(15) = 2.49, P = 0.025; R^2 = 0.36$; right STSp_upper: $b = 8.42, 95\% \text{ CI} = 0.49\text{--}16.35, t(15) = 2.13, P = 0.05; R^2 = 0.31$. None of these results, however, survived the FDR correction across the thirty comparisons, though notably none of the regressions for the No Gesture+Specific Language condition were significant, even before correction, and even in the STSp_upper bank: left STSp_upper: $b = 0.74, 95\% \text{ CI} = -4.20 \text{ to } 5.79, t(15) = 0.30, P = 0.77; R^2 = 0.002$, right STSp_upper: $b = 2.75, 95\% \text{ CI} = -1.40 \text{ to } 7.84, t(15) =$

1.21, $P = 0.25; R^2 = 0.04$. This pattern suggests that the modulation of activity in the upper bank of the STSp might be related to gesture processing rather than to language processing more generally, but because no finding survived the statistical correction, we must be cautious in this interpretation.

DISCUSSION

Our goal was to characterize the brain regions that are involved in integrating semantic information from gesture and speech, with a particular focus on the role that inferior frontal and posterior temporal brain regions play in this

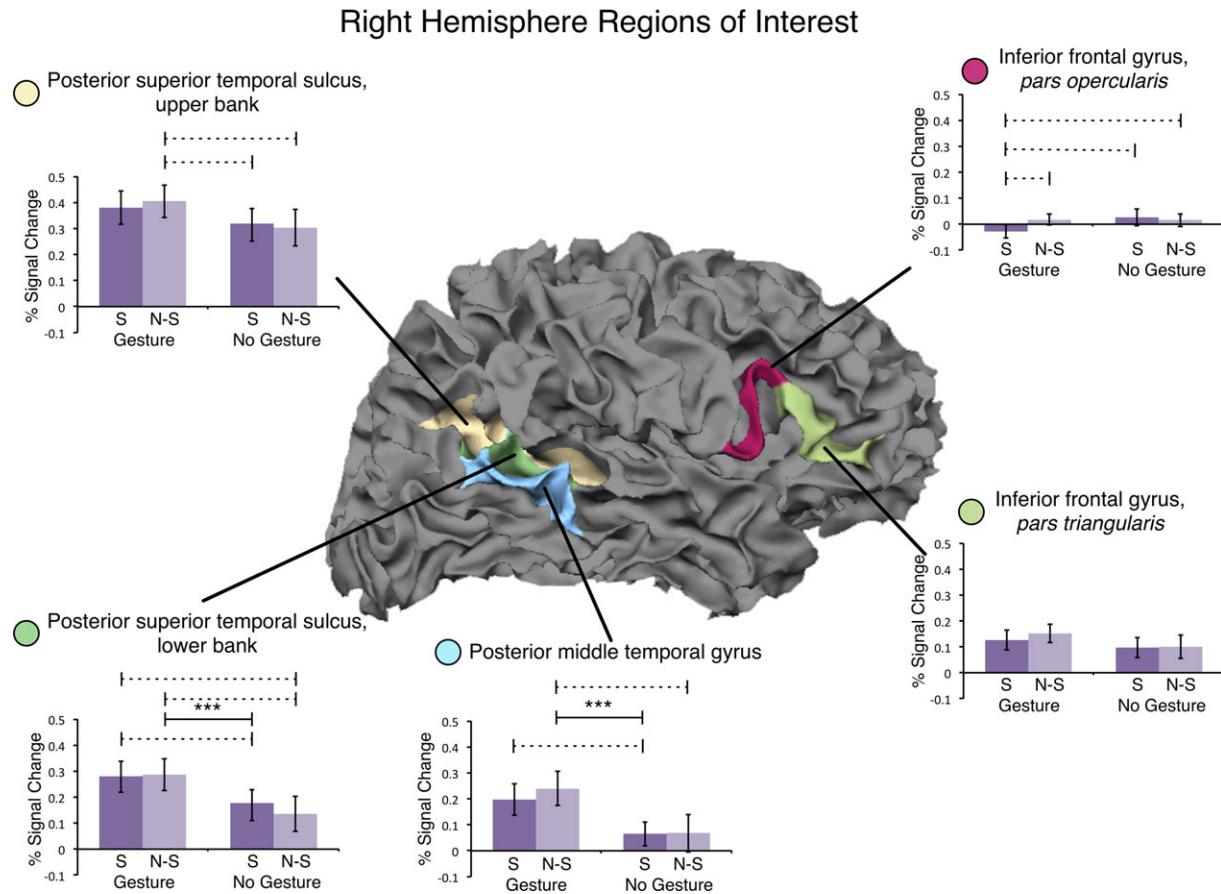


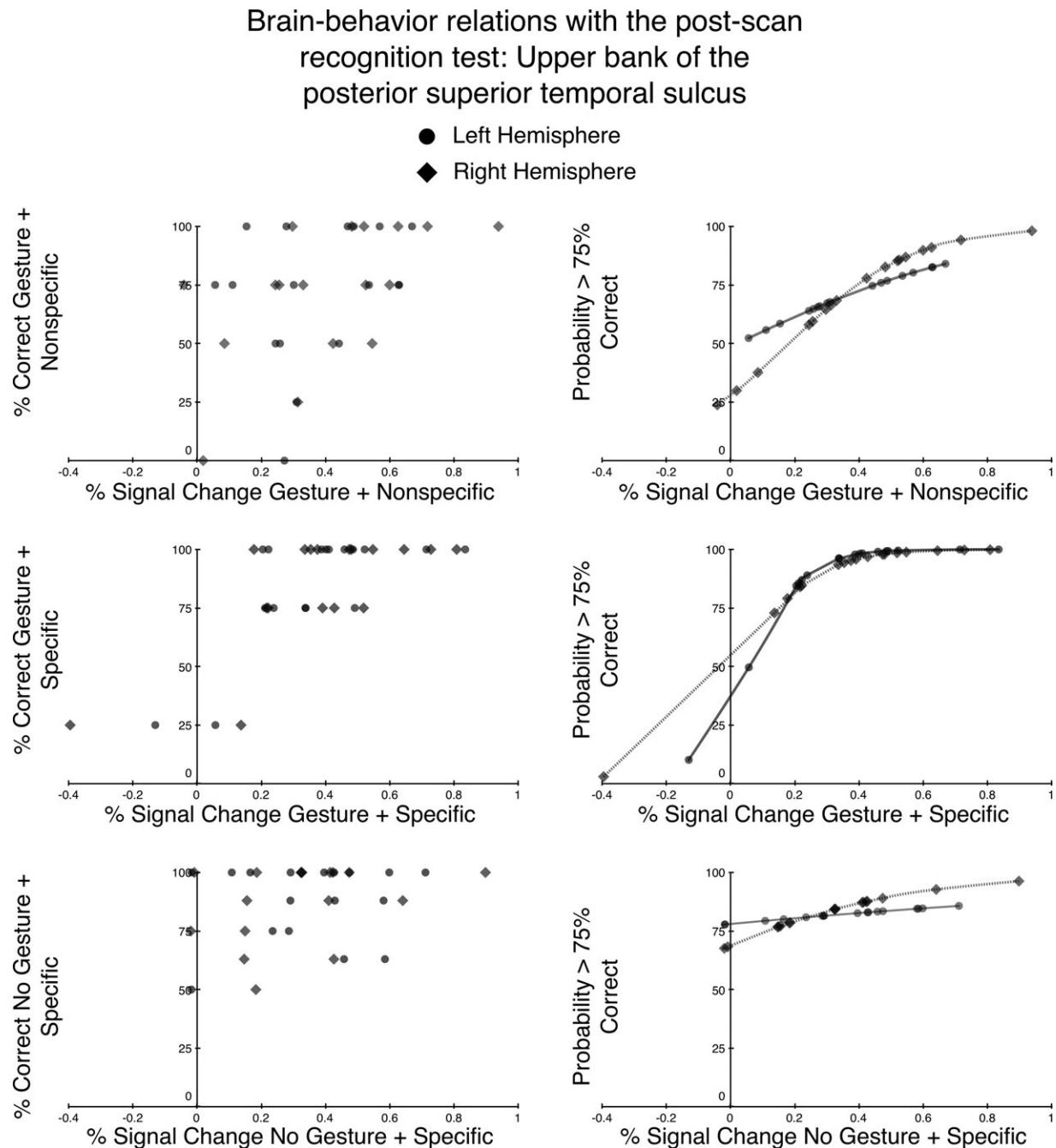
Figure 5.

Results of the analysis for right hemisphere inferior frontal and posterior temporal regions of interest. Regions are outlined on a white matter surface image of a representative subject with the pial surface removed to reveal the sulci. In the right hemisphere, we found no significant interactions across the Language Specificity and Gesture factors. S, specific language; N-S, nonspecific language. *** $P < 0.001$ (corrected); ** $P < 0.01$ (corrected); --- $P < 0.05$ (uncorrected).

process. We found that three regions—left IFGTr, left IFGOP, and left MTGp—responded more strongly when the speaker’s iconic gestures added to the information in speech (Gesture+Nonspecific Language) than when the same gestures reinforced the information in speech (Gesture+Specific Language). We suggest that the increased activation in these areas is reflecting the cognitive demand required to integrate the additional information conveyed in gesture with the information conveyed in speech into a coherent semantic representation. Importantly, however, when the same speech was not accompanied by gesture, these regions did not respond more strongly to Nonspecific vs. Specific Language, suggesting that the areas are responding specifically to the task of integrating information across gesture and speech (and not to the task of interpreting nonspecific speech). These results demonstrate that IFGTr, IFGOP, and MTGp are not only sensitive to the meaning conveyed by iconic gestures, but also to the

relation between the information conveyed in gesture and speech.

We also found that activity in the left STSp, previously implicated in gesture-speech integration [Holle et al., 2008, 2010], was not sensitive to the gesture-speech interaction (i.e., it was not sensitive to the meaning of the gestures). However, during the processing of stories with iconic gestures, activity in this region was correlated with recall of details of the stories in a post-scan recognition test, while no relation was found for the recall of stories without gestures. Notably, though, these correlations did not survive a statistical correction. Thus, this provides only weak evidence for the role of STSp within a broader frontal-temporal-parietal network for gesture-speech integration, possibly by enhancing attention to gesture and speech during naturalistic audiovisual situations. Together, these findings clarify the neurobiology of gesture-speech integration and contribute to an emerging picture of how listeners glean meaning from the gestures that accompany speech.

**Figure 6.**

Results of proportional odds logistic regression analysis showing the relationship between fMRI signal in both left and right upper banks of the posterior superior temporal sulcus and recall of the details of the Gesture+Nonspecific (top), the Gesture+Specific (middle), and No Gesture+Specific (bottom) stories. Left

side of figure shows the plot of the raw data. Right side of figure shows the probability of performing better than 75% on the post-scan recognition test, predicted by activity in the upper bank of posterior the superior temporal sulcus, left hemisphere (circles), and right hemisphere (diamonds).

Inferior Frontal Gyrus and Posterior Middle Temporal Gyrus Contribute to Gesture-Speech Integration

We found that both IFGTr and IFGOp and left MTGp responded more strongly to nonspecific than to specific language, but only when the nonspecific language was accompanied by a gesture that narrowed down the range of possible interpretations of the language. Thus, left IFG and MTGp play a role in connecting meaning from gesture and speech, that is, in gesture-speech integration. In contrast to previous work claiming that semantic integration occurs in either left IFG [Willems et al., 2007] or more posterior temporal regions of the left hemisphere [Holle et al., 2008, 2010], our results suggest that both left IFG and MTGp regions contribute to this process. Our finding is consistent with a significant body of empirical research that has emphasized the contribution of IFG and MTGp to processing semantic information during language comprehension without gesture [Kutas and Federmeier, 2011; Lau et al., 2008; Van Petten and Luka, 2006; Vigneau et al., 2006; for review]. Both regions have been associated with semantic activation and meaning selection within a broader semantic context, with the IFG particularly involved in controlled retrieval and selection among competing semantic representations [Badre and Wagner, 2007; Badre et al., 2005; Devlin et al., 2003; Gough et al., 2005; Jefferies et al., 2008; Moss et al., 2005; Thompson-Schill et al., 1997; Wagner et al., 2001; Ye and Zhou, 2009].

Research consistently shows that brain activity in left IFG increases when participants are required to retrieve the nondominant or subordinate meanings of ambiguous words or sentences, which increases demand on semantic selection and retrieval [Fig. 1; Bedny et al., 2008; Gennari et al., 2007; Hoenig and Scheef, 2009; Rodd et al., 2005; Whitney et al., 2009; Zempleni et al., 2007]. Further, left IFG activity and connectivity increase when iconic gestures are incongruent with or unrelated to the accompanying speech, although this process tends to be associated with more anterior portions of IFG [Green et al., 2009; Skipper et al., 2007; Willems et al., 2007, 2009]. Notably, we found that both the anterior (IFGTr) and the posterior (IFGOp) IFG contributed to the semantic integration process. This finding is somewhat inconsistent with previous research emphasizing the contribution of posterior IFG to phonological, rather than semantic processing [Gold et al., 2006; Gough et al., 2005; Hickok and Poeppel, 2007], although some language studies do report activity peaks in the posterior IFG in response to demands on semantic selection and retrieval [see Fig. 1; Bedny et al., 2008; Sniijders et al., 2009; Zempleni et al., 2007; Lau et al., 2008 for review]. Further, in a previous study, we found that both anterior and posterior regions of the right IFG were sensitive to the relation between gesture and speech [Dick et al., 2009], indicating that the contribution of posterior IFG to semantic processing cannot be ruled out. Moreover, its involvement in semantic processing is consistent with the

proposal by Lau et al. [2008] that posterior IFG mediates selection between highly activated candidate representations, and anterior IFG mediates controlled retrieval of semantic information. Both processes are potentially involved in gesture-speech integration.

Left MTGp is also considered to be a critical component of a semantic network, and thought to be involved in the long-term storage of semantic information [Binder et al., 2009; Binney et al., 2010; Doehrmann and Naumer, 2008; Hickok and Poeppel, 2007; Martin and Chao, 2001; Price, 2010; Rogers et al., 2004; Vandenberghe et al., 1996]. However, left MTGp tends to coactivate with left IFG in response to increased demands on semantic selection and retrieval [see Fig. 1; Badre et al., 2005; Gold et al., 2006; Hoenig and Scheef, 2009; Kuperberg et al., 2008; Thompson-Schill et al., 1997; Wagner et al., 2001; Zempleni et al., 2007] and, more recently, left MTGp has been implicated in controlled retrieval of semantic information [Whitney et al., 2011a,b]. Whitney et al. [2011b] investigated this possibility more closely using repetitive transcranial magnetic stimulation (rTMS), which they applied to both left IFG and left MTGp during a semantic judgment task. Here, the primary manipulation was whether a cue was strongly or weakly related to a target. rTMS interfered with responding during the weakly cued, but not the strongly cued, "automatic" task, which was taken to suggest that the weak cue requires controlled semantic retrieval processes not needed in the strong cue. Interestingly, this interference occurred for rTMS to both left IFG and to left MTGp, suggesting a role in controlled retrieval for both regions.

Our findings do not fall strongly on either side of this debate—that is, they do not provide strong support for MTGp as a region involved in either controlled retrieval of activated representations, or the storage of those representations. Although we did find that left MTGp was sensitive to the semantic manipulation—it responds more strongly to nonspecific language, but only when iconic gestures disambiguate the meaning of that language—such a finding is consistent with both semantic activation and semantic selection accounts. In fact, we suggest that, to some degree, this is a false dichotomy. There is a division of labor between IFG and MTGp in processing semantic information, suggested in this study by the different patterns of activity in these regions in response to language without gesture (Figs. 3 and 4). However, retrieval and selection processes require activation of stored representations, which would lead to the activation of both regions. It may, therefore, be more fruitful to hypothesize a collaboration between IFG and MTGp in retrieving and manipulating semantic knowledge stored in a distributed fashion in other parts of the temporal and parietal cortices [cf. Whitney et al., 2011a]. In other words, the notion that a single region "does" semantic integration of gesture and speech [Holle et al., 2008, 2010; Willems et al., 2007] is both misguided and inconsistent with the empirical picture emerging from functional imaging studies of gesture [e.g., see Willems et al., 2009].

Posterior Superior Temporal Sulcus Contributions to Gesture-Speech Integration

We failed to find evidence that STSp—either the upper or lower bank—was involved in processing iconic gestures with speech at the semantic level. This finding is consistent with several previous studies of gesture-speech integration [Willems et al., 2007, 2009], including work from our own lab [Dick et al., 2009, 2012; Skipper et al., 2009]. In the past [Dick et al., 2009, 2012], we have attributed activity in the STSp during the processing of co-speech gesture to this region's putative role in processing biologically relevant motion from hand movements [Beauchamp et al., 2003; Grossman et al., 2000; Saygin et al., 2004; Thompson et al., 2005]. However, two studies [Holle et al., 2008; Straube et al., 2011] have reported sensitivity in the STSp to semantic information in gesture. Further, in the present study, we found only weak evidence (Fig. 6) that activity in the upper bank of both right and left STSp was correlated with recall of stories that contained iconic gestures. Broadly, our findings are in agreement with the proposal put forth by Holle et al. [2010] that attending to gesture can enhance speech comprehension, particularly under adverse listening conditions, and that this enhancement recruits the bilateral STSp.

However, we disagree with one key aspect of Holle et al. [2010] proposal. As already discussed, we do not subscribe to the idea that “integration of iconic gestures and speech takes place at the posterior end of the superior temporal sulcus and adjacent superior temporal gyrus (pSTS/STG)” [Holle et al., 2010, p. 882]. We believe that the STSp participates with other inferior frontal, temporal, and inferior parietal brain regions to accomplish gesture-speech integration. Indeed, in a recent developmental study [Dick et al., 2012], we found that gesture meaning failed to modulate BOLD signal amplitude in STSp. However, using structural equation modeling of effective connectivity among frontal, temporal, and parietal brain regions, we showed that the age difference in the strength of connectivity between STSp and other temporal and inferior parietal brain regions was moderated by the semantic relation between gesture and speech. That is, for adults, effective connectivity among these regions was greater when viewing meaningful gestures compared to nonmeaningful gestures, but effective connectivity across these two conditions did not differ for children. Considered in light of the present findings and those of Holle et al. [2010], the Dick et al. [2012] findings suggest that STSp may be involved in connecting information from the visual and auditory modalities, and through its interactions with other brain regions participates in constructing a coherent meaning from gesture and speech. This idea is also consistent with the known connectivity of the upper bank of the STSp, which receives a predominance of afferents from auditory and visual association areas of the superior temporal, inferior parietal, and occipital cortices involved in processing information in the auditory and visual modalities [Seltzer and Pandya, 1978, 1989, 1994].

Left and Right Hemisphere Contributions to Gesture-Speech Integration

Although there is a bias to focus on left hemisphere contributions to language and, by extension, gesture comprehension, it is clear that the right hemisphere also contributes to both processes, particularly during narrative-level language comprehension [Ferstl et al., 2008; Jung-Beeman, 2005; Dick et al., 2009; Wilson et al., 2008]. Further, in some instances right IFG and MTGp regions also show sensitivity to increased demand on semantic retrieval and selection during language comprehension without gesture [Hein et al., 2007; Hoenig and Scheef, 2009; Lauro et al., 2008; Rodd et al., 2005; Snijders et al., 2009; Stowe et al., 2005; Zempleni et al., 2007]. In addition, right IFG activity is associated with the correct recall of non-meaningful hand movements that accompany speech [Straube et al., 2009], and with the meaningfulness of gestures that accompany speech [Dick et al., 2009; 2011; Green et al., 2009]. Despite these findings, our anatomical ROI analysis revealed no brain regions on the right hemisphere to be sensitive to the semantic relation between gesture and speech. The conflict between these findings and prior research that has found right hemisphere sensitivity to gesture semantics can be explained by the nature of the hand movements accompanying speech. In the previous studies, the right IFG responded more strongly when the hand movement was unrelated in a clear way to the speech (e.g., the hand movement was a grooming movement), and thus required additional effort to fit the gesture to the content of the accompanying speech. In this study, the iconic gesture was always interpretable in the context of speech. Under this view, right hemisphere recruitment for gesture-speech integration, particularly for the right IFG, requires situations in which the demand on semantic retrieval and selection is very high [Chou et al., 2006; Hein et al., 2007; Rodd et al., 2005]. When the gestures are relatively easy to integrate with speech—that is, when they are iconic and have a meaning that is relatively transparent—the demand is lower and recruits only the left hemisphere regions [Willems et al., 2007, 2009].

SUMMARY

The findings reported here help to clarify and elaborate an emerging picture of how the brain integrates information conveyed in gesture and speech. We have found that inferior frontal and posterior middle temporal cortical regions, particularly in the left hemisphere, are directly involved in constructing a unitary semantic interpretation from separate auditory (speech) and visual (gesture) modalities. The STSp also appears to play a role, but it may be more heavily involved in directing attention to gestures, rather than integrating information across gesture and speech. This picture is consistent with studies of language comprehension without gesture, which find that

inferior frontal and posterior middle temporal regions play an important role in the activation and retrieval of semantic information in general.

REFERENCES

- Alibali MW, Goldin-Meadow S (1993): Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cogn Psychol* 25: 468–468.
- Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings HB, Zilles K (1999): Broca's region revisited: Cytoarchitecture and intersubject variability. *J Comp Neurol* 412: 319–341.
- Argall BD, Saad ZS, Beauchamp MS (2006): Simplified intersubject averaging on the cortical surface using SUMA. *Hum Brain Mapp* 27: 14–27.
- Badre D, Poldrack RA, Pare-Blaauw EJ, Insler RZ, Wagner AD (2005): Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* 47: 907–918.
- Badre D, Wagner AD (2007): Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia* 45: 2883–2901.
- Bates E, Dick F (2002): Language, gesture, and the developing brain. *Dev Psychobiol* 40: 293–310.
- Beattie G, Shovelton H (1999): Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *J Lang Soc Psychol* 18: 438–462.
- Beattie G, Shovelton H (2002): An experimental investigation of some properties of individual iconic gestures that mediate their communicative power. *Br J Psychol* 93: 179–192.
- Beauchamp MS, Lee KE, Haxby JV, Martin A (2003): fMRI responses to video and point-light displays of moving humans and manipulable objects. *J Cogn Neurosci* 15: 991–1001.
- Bedny M, McGill M, Thompson-Schill SL (2008): Semantic adaptation and competition during word comprehension. *Cereb Cortex* 18: 2574–2585.
- Bedny M, McGill M, Thompson-Schill SL (2008): Semantic adaptation and competition during word comprehension. *Cereb Cortex* 18: 2574–2585.
- Benjamini Y, Gavrilov Y (2009): A simple forward selection procedure based on false discovery rate control. *Annals Appl Stat* 3: 179–198.
- Benjamini Y, Hochberg Y (1995): Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Roy Stat Soc: Series B (Stat Methodol)* 57: 289–300.
- Benjamini Y, Hochberg Y (1995): Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Roy Stat Soc: Series B (Stat Methodol)* 57: 289–300.
- Benjamini Y, Hochberg Y (2000): On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J Educ Behav Stat* 25: 60.
- Binder JR, Desai RH, Graves WW, Conant LL (2009): Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex* 19: 2767–2796.
- Binney RJ, Embleton KV, Jefferies E, Parker GJ, Ralph MA (2010): The ventral and inferolateral aspects of the anterior temporal lobe are crucial in semantic memory: Evidence from a novel direct comparison of distortion-corrected fMRI, rtMS, and semantic dementia. *Cereb Cortex* 20: 2728–2738.
- Chou TL, Booth JR, Burman DD, Bitan T, Bigio JD, Lu D, Cone NE (2006): Developmental changes in the neural correlates of semantic processing. *Neuroimage* 29: 1141–1149.
- Chung MK, Robbins SM, Dalton KM, Davidson RJ, Alexander AL, Evans AC (2005): Cortical thickness analysis in autism with heat kernel smoothing. *Neuroimage* 25: 1256–1265.
- Church RB, Goldin-Meadow S (1986): The mismatch between gesture and speech as an index of transitional knowledge. *Cognition* 23: 43–71.
- Cox RW (1996): AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29: 162–173.
- Dale AM, Fischl B, Sereno MI (1999): Cortical surface-based analysis I: Segmentation and surface reconstruction. *Neuroimage* 9: 179–194.
- Desai R, Liebenthal E, Possing ET, Waldron E, Binder JR (2005): Volumetric vs. surface-based alignment for localization of auditory cortex activation. *Neuroimage* 26: 1019–1029.
- Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ (2006): An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31: 968–980.
- Devlin JT, Matthews PM, Rushworth MFS (2003): Semantic processing in the left inferior prefrontal cortex: A combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience* 15: 71–84.
- Devlin JT, Poldrack RA (2007): In praise of tedious anatomy. *Neuroimage* 37: 1033–1041.
- Dick AS, Goldin-Meadow S, Hasson U, Skipper JL, Small SL (2009): Co-Speech gestures influence neural activity in brain regions associated with processing semantic information. *Hum Brain Mapp* 30: 3509–3526.
- Dick AS, Goldin-Meadow S, Solodkin A, Small SL (2012): Gesture in the developing brain. *Dev Sci* 15:165–180. DOI: 10.1111/j.1467-7687.2011.01100.x
- Doehrmann O, Naumer MJ (2008): Semantics and the multisensory brain: How meaning modulates processes of audio-visual integration. *Brain Res* 1242: 136–150.
- Duvernoy HM (1999): The Human Brain: Surface, Blood Supply, and Three-Dimensional Sectional Anatomy. New York:Springer-Verlag.
- Ferstl EC, Neumann J, Bogler C, Yves von Cramon D (2008): The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Hum Brain Mapp* 29:581–593.
- Feyereisen P (2006): Further investigation on the mnemonic effect of gestures: Their meaning matters. *Eur J Cogn Psychol* 18: 185–205.
- Feyereisen P, Van de Wiele M, Dubois F (1988): The meaning of gestures: What can be understood without speech? *Cahiers de Psychologie Cognitive* 8: 3–25.
- Fischl B, Sereno MI, Dale AM (1999): Cortical surface-based analysis. II. Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9: 195–207.
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM (2004): Automatically parcellating the human cerebral cortex. *Cereb Cortex* 14: 11–22.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995): Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magn Resonance Med* 33: 636–647.
- Fox J (2002): An R and s-plus companion to applied regression. Sage Publications, Inc.
- Gennari SP, MacDonald MC, Postle BR, Seidenberg MS (2007): Context-dependent interpretation of words: Evidence for interactive neural processes. *Neuroimage* 35: 1278–1286.

- Genovese CR, Lazar NA, Nichols T (2002): Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15: 870–878.
- Gold BT, Balota DA, Jones SJ, Powell DK, Smith CD, Andersen AH (2006): Dissociation of automatic and strategic lexical semantics: Functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *J Neurosci* 26: 6523–6532.
- Goldin-Meadow S (2003): Hearing Gesture: How Our Hands Help Us Think. Cambridge, MA: Harvard University Press (Belknap Division).
- Gough PM, Nobre AC, Devlin JT (2005): Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. *J Neurosci* 25: 8010–8016.
- Green A, Straube B, Weis S, Jansen A, Willmes K, Konrad K, Kircher T (2009): Neural integration of iconic and unrelated coverbal gestures: A functional MRI study. *Hum Brain Mapp* 30: 3309–3324.
- Grossman E, Donnelly M, Price R, Pickens D, Morgan V, Neighbor G, Blake R (2000): Brain areas involved in perception of biological motion. *J Cogn Neurosci* 12: 711–720.
- Habets B, Kita S, Shao Z, Özyürek A, Hagoort P (2011): The role of synchrony and ambiguity in speech-gesture integration during comprehension. *J Cogn Neurosci* 23: 1845–1854.
- Hardin JW, Hilbe J (2007): Generalized Linear Models and Extensions. College Station, TX:Stata Corp.
- Hein G, Doehrmann O, Müller NG, Kaiser J, Muckli L, Naumer MJ (2007): Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. *J Neurosci* 27: 7881–7887.
- Hickok G, Poeppel D (2007): The cortical organization of speech processing. *Nat Rev Neurosci* 8: 393–402.
- Hoenig K, Scheef L (2009): Neural correlates of semantic ambiguity processing during context verification. *Neuroimage* 45: 1009–1019.
- Holle H, Gunter TC (2007): The role of iconic gestures in speech disambiguation: ERP evidence. *J Cogn Neurosci* 19: 1175–1192.
- Holle H, Gunter TC, Rüschemeyer SA, Hennenlotter A, Iacoboni M (2008): Neural correlates of the processing of co-speech gestures. *Neuroimage* 39: 2010–2024.
- Holle H, Obleser J, Rüschemeyer S-A, Gunter TC (2010): Integration of iconic gestures and speech in left superior temporal areas boosts speech comprehension under adverse listening conditions. *Neuroimage* 49: 875–884.
- Holler J, Beattie G (2003): Pragmatic aspects of representational gestures: Do speakers use them to clarify verbal ambiguity for the listener? *Gesture* 3: 127–154.
- Jefferies E, Patterson K, Ralph MA (2008): Deficits of knowledge versus executive control in semantic cognition: Insights from cued naming. *Neuropsychologia* 46: 649–658.
- Jung-Beeman M (2005): Bilateral brain processes for comprehending natural language. *Trends Cogn Sci* 9: 712–718.
- Kelly SD, Barr DJ, Church RB, Lynch K (1999): Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *J Memory Lang* 40: 577–592.
- Kelly SD, Church RB (1998): A comparison between children's and adults' ability to detect conceptual information conveyed through representational gestures. *Child Dev* 69: 85–93.
- Kelly SD, Creigh P, Bartolotti J (2010): Integrating speech and iconic gestures in a stroop-like task: Evidence for automatic processing. *J Cogn Neurosci* 22: 683–694.
- Kelly SD, Kravitz C, Hopkins M (2004): Neural correlates of bimodal speech and gesture comprehension. *Brain Lang* 89: 253–260.
- Kircher T, Straube B, Leube D, Weis S, Sachs O, Willmes K, Konrad K, Green A (2009): Neural interaction of speech and gesture: Differential activations of metaphoric co-verbal gestures. *Neuropsychologia* 47: 169–179.
- Krishnamoorthy K, Lu F, Mathew T (2007): A parametric bootstrap approach for ANOVA with unequal variances: Fixed and random models. *Computat Stat Data Anal* 51: 5731–5742.
- Kuperberg GR, Sitnikova T, Lakshmanan BM (2008): Neuroanatomical distinctions within the semantic system during sentence comprehension: Evidence from functional magnetic resonance imaging. *Neuroimage* 40: 367–388.
- Kutas M, Federmeier KD (2011): Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annu Rev Psychol* 62: 621–647.
- Lau EF, Phillips C, Poeppel D (2008): A cortical network for semantics:(De) constructing the N400. *Nat Rev Neurosci* 9: 920–933.
- Lauro LJ, Tettamanti M, Cappa SF, Papagno C (2008): Idiom comprehension: A prefrontal task? *Cereb Cortex* 18: 162–170.
- Mai JK, Paxinos G, Voss T 2007. *Atlas of the Human Brain*. Academic Press:San Diego.
- Malikovic A, Amunts K, Schleicher A, Mohlberg H, Eickhoff SB, Wilms M, Palomero-Gallagher N, Armstrong E, Zilles K (2007): Cytoarchitectonic analysis of the human extrastriate cortex in the region of V5/MT: A probabilistic, stereotaxic map of area hOc5. *Cereb Cortex* 17: 562–574.
- Mar RA (2011): The neural bases of social cognition and story comprehension. *Annu Rev Psychol* 62: 103–134.
- Martin A, Chao LL (2001): Semantic memory and the brain: Structure and processes. *Curr Opin Neurobiol* 11: 194–201.
- McNeill D 1992. *Hand and mind: What gestures reveal about thought*.Chicago:University of Chicago Press.
- McNeill D, Cassell J, McCullough K-E (1994): Communicative effects of speech-mismatched gestures. *Res Lang Soc Interact* 27: 223–237.
- Moss HE, Abdallah S, Fletcher P, Bright P, Pilgrim L, Acres K, Tyler LK (2005): Selecting among competing alternatives: Selection and retrieval in the left inferior frontal gyrus. *Cereb Cortex* 15: 1723–1735.
- Neito-Castanon,A, Ghosh,SS, Tourville,JA, &Guenther,FH (2003): Region of interest based analysis of functional imaging data. *Neuroimage* 19: 1303–1316.
- Nichols T, Brett M, Andersson J, Wager T, Poline JB (2005): Valid conjunction inference with the minimum statistic. *Neuroimage* 25: 653–660.
- Nichols T, Brett M, Andersson J, Wager T, Poline JB (2005): Valid conjunction inference with the minimum statistic. *Neuroimage* 25: 653–660.
- Özyürek A, Willem RM, Kita S, Hagoort P (2007): On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials. *J Cogn Neurosci* 19: 605–616.
- Perry M, Church RB, Goldin-Meadow S (1992): Is gesture-speech mismatch a general index of transitional knowledge?* *I. Cogn Dev* 7: 109–122.
- Price CJ (2010): The anatomy of language: A review of 100 fmri studies published in 2009. *Ann NY Acad Sci* 1191: 62–88.
- Rodd JM, Davis MH, Johnsrude IS (2005): The neural mechanisms of speech comprehension: FMRI studies of semantic ambiguity. *Cerebral Cortex* 15: 1261–1269.
- Rogers TT, Lambon Ralph MA, Garrard P, Bozeat S, McClelland JL, Hodges JR, Patterson K (2004): Structure and deterioration

- of semantic memory: A neuropsychological and computational investigation. *Psychol Rev* 111: 205–235.
- Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry M-S, Umarova R, Musso M, Glauche V, Abel S, Huber W, Rijntjes M, Hennig J, Weiller C (2008): Ventral and dorsal pathways for language. *Proc Natl Acad Sci USA* 105: 18035–18040.
- Saygin AP, Wilson SM, Hagler DJ Jr, Bates E, Sereno MI (2004): Point-Light biological motion perception activates human pre-motor cortex. *J Neurosci* 24: 6181–6188.
- Schmahmann JD, Pandya DN (2006): *Fiber Pathways of the Brain*. Oxford, England: Oxford University Press.
- Seltzer B, Pandya DN (1978): Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Res* 149: 1–24.
- Seltzer B, Pandya DN (1989): Intrinsic connections and architectonics of the superior temporal sulcus in the rhesus monkey. *J Comp Neurol* 290: 451–471.
- Seltzer B, Pandya DN (1994): Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: A retrograde tracer study. *J Comparative Neurol* 343: 445–463.
- Skipper JI, Goldin-Meadow S, Nusbaum HC, Small SL (2007): Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language* 101: 260–277.
- Skipper JI, Goldin-Meadow S, Nusbaum HC, Small SL (2009): Gestures orchestrate brain networks for language understanding. *Current Biology* 19: 1–7.
- Skipper JI, Nusbaum HC, Small SL (2005): Listening to talking faces: Motor cortical activation during speech perception. *Neuroimage* 25: 76–89.
- Small SL, Nusbaum HC (2004): On the neurobiological investigation of language understanding in context. *Brain Lang* 89: 300–311.
- Snijders TM, Vosse T, Kempen G, Van Berkum JJA, Petersson KM, Hagoort P (2009): Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word-category ambiguity. *Cerebral Cortex* 19: 1493–1503.
- Stowe LA, Haverkort M, Zwarts F (2005): Rethinking the neurological basis of language. *Lingua* 115: 997–1042.
- Straube B, Green A, Bromberger B, Kircher T (2011): The differentiation of iconic and metaphoric gestures: Common and unique integration processes. *Hum Brain Mapp* 32: 520–533.
- Straube B, Green A, Jansen A, Chatterjee A, Kircher T (2010): Social cues, mentalizing and the neural processing of speech accompanied by gestures. *Neuropsychologia* 48: 382–393.
- Straube B, Green A, Weis S, Chatterjee A, Kircher T (2009): Memory effects of speech and gesture binding: Cortical and hippocampal activation in relation to subsequent memory performance. *J Cogn Neurosci* 21: 821–836.
- Thompson JC, Clarke M, Stewart T, Puce A (2005): Configural processing of biological motion in human superior temporal sulcus. *J Neurosci* 25: 9059–9066.
- Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ (1997): Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proc Natl Acad Sci USA* 94: 14792–14797.
- Vandenbergh R, Price C, Wise R, Josephs O, Frackowiak RS (1996): Functional anatomy of a common semantic system for words and pictures. *Nature* 383: 254–256.
- Van Essen DC, Dierker D, Snyder AZ, Raichle ME, Reiss AL, Korenberg J (2006): Symmetry of cortical folding abnormalities in Williams syndrome revealed by surface-based analyses. *J Neurosci* 26: 5470–5483.
- Van Petten C, Luka BJ (2006): Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain Lang* 97: 279–293.
- Venables WN, Ripley BD (2010): *Modern Applied Statistics with S*. Springer-Verlag.
- Vigneau M, Beaucousin V, Hervé PY, Duffau H, Crivello F, Houdé O, Mazoyer B, Tzourio-Mazoyer N (2006): Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage* 30: 1414–1432.
- Wagner AD, Paré-Blairev EJ, Clark J, Poldrack RA (2001): Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron* 31: 329–338.
- Whitney C, Grossman M, Kircher TT (2009): The influence of multiple primes on bottom-up and top-down regulation during meaning retrieval: Evidence for 2 distinct neural networks. *Cereb Cortex* 19: 2548–2560.
- Whitney C, Jefferies E, Kircher T (2011): Heterogeneity of the left temporal lobe in semantic representation and control: Priming multiple versus single meanings of ambiguous words. *Cereb Cortex* 21: 831–844.
- Whitney C, Kirk M, O'Sullivan J, Lambon Ralph MA, Jefferies E (2011): The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cereb Cortex* 21: 1066–1075.
- Wilcox RR (2012): *Introduction to Robust Estimation and Hypothesis Testing*, 3rd Ed. Waltham, MA: Academic Press.
- Wilkinson L (1999): Statistical methods in psychology journals: Guidelines and explanations. *Am Psychologist* 54: 594.
- Willems RM, Özyürek A, Hagoort P (2007): When language meets action: The neural integration of gesture and speech. *Cereb Cortex* 17: 2322.
- Willems RM, Özyürek A, Hagoort P (2009): Differential roles for left inferior frontal and superior temporal cortex in multimodal integration of action and language. *Neuroimage* 47: 1992–2004.
- Wilson SM, Molnar-Szakacs I, Iacoboni M (2008): Beyond superior temporal cortex: Intersubject correlations in narrative speech comprehension. *Cereb Cortex* 18: 230–242.
- Wright DB, London K (2009): *Modern Regression Techniques Using R: A Practical Guide*. London: Sage Publications Ltd.
- Wu YC, Coulson S (2007a): How iconic gestures enhance communication: An ERP study. *Brain Lang* 101: 234–245.
- Wu YC, Coulson S (2007b): Iconic gestures prime related concepts: An ERP study. *Psychonomic Bull Rev* 14: 57–63.
- Ye Z, Zhou X (2009): Executive control in language processing. *Neurosci Biobehav Rev* 33: 1168–1177.
- Zempleni MZ, Renken R, Hoeks JC, Hoogduin JM, Stowe LA (2007): Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *Neuroimage* 34: 1270–1279.
- Zeno S, Associates TAS, Health NIOC, (US) HD, Science NYS, Foundation T (1995): *The Educator's Word Frequency Guide*. NJ: Touchstone Applied Science Associates Brewster.