ELSEVIER

Speech-associated gestures, Broca's area, and the human mirror system

Jeremy I. Skipper [a,b,*], Susan Goldin-Meadow [a], Howard C. Nusbaum [a], Steven L. Small [a,b]

[a] *Department of Psychology, The University of Chicago, USA*
[b] *Department of Neurology, MC 2030, and the Brain Research Imaging Center, The University of Chicago, 5841 South Maryland Ave., Chicago, IL 60637, USA*

Abstract

Speech-associated gestures are hand and arm movements that not only convey semantic information to listeners but are themselves actions. Broca's area has been assumed to play an important role both in semantic retrieval or selection (as part of a language comprehension system) and in action recognition (as part of a "mirror" or "observation–execution matching" system). We asked whether the role that Broca's area plays in processing speech-associated gestures is consistent with the semantic retrieval/selection account (predicting relatively weak interactions between Broca's area and other cortical areas because the meaningful information that speech-associated gestures convey reduces semantic ambiguity and thus reduces the need for semantic retrieval/selection) or the action recognition account (predicting strong interactions between Broca's area and other cortical areas because speech-associated gestures are goal-direct actions that are "mirrored"). We compared the functional connectivity of Broca's area with other cortical areas when participants listened to stories while watching meaningful speech-associated gestures, speech-irrelevant self-grooming hand movements, or no hand movements. A network analysis of neuroimaging data showed that interactions involving Broca's area and other cortical areas were *weakest* when spoken language was accompanied by meaningful speech-associated gestures, and *strongest* when spoken language was accompanied by self-grooming hand movements or by no hand movements at all. Results are discussed with respect to the role that the human mirror system plays in processing speech-associated movements.
© 2007 Published by Elsevier Inc.

Keywords: Language; Gesture; Face; The motor system; Premotor cortex; Broca's area; Pars opercularis; Pars triangularis; Mirror neurons; The human mirror system; Action recognition; Action understanding; Structural equation models

1. Introduction

Among the actions that we encounter most in our lives are those that accompany speech during face-to-face communication. Speakers often move their hands when they talk (even when a listener cannot see the speaker's hand, Rimé, 1982). These hand movements, called speech-associated gestures, are distinct from codified emblems (e.g., "thumbs-up"), pantomime, and sign language in their reliance on, and co-occurrence with, spoken language (McNeill, 1992).

Speech-associated gestures often convey information that complements the information conveyed in the talk they accompany and, in this sense, are meaningful (Goldin-Meadow, 2003). For this reason, such hand and arm actions have variously been called "representational gestures" (McNeill, 1992), "illustrators" (Ekman & Friesen, 1969), "gesticulations" (Kendon, 2004), and "lexical gestures" (Krauss, Chen, & Gottesman, 2000). Consistent with the claim that speech-associated gestures convey information that complements the information conveyed in talk, speech-associated gestures have been found to improve listener comprehension, suggesting that they are meaningful to listeners (Alibali, Flevares, & Goldin-Meadow, 1997; Berger & Popelka, 1971; Cassell, McNeill, & McCullough, 1999; Driskell & Radtke, 2003; Goldin-Meadow & Momeni Sandhofer, 1999; Goldin-Meadow, Wein,

---

* Corresponding author. Address: Department of Neurology, MC 2030, and the Brain Research Imaging Center, The University of Chicago, 5841 South Maryland Ave., Chicago, IL 60637, USA. Fax: +1 773 834 7610.
 E-mail address: skipper@uchicago.edu (J.I. Skipper).

& Chang, 1992; Kendon, 1987; McNeill, Cassell, & McCullough, 1994; Records, 1994; Riseborough, 1981; Rogers, 1978; Singer & Goldin-Meadow, 2005; Thompson & Massaro, 1986). Speech-associated gestures are thus hand movements that provide accessible semantic information relevant to language comprehension.

Broca's area has been implicated in spoken language comprehension, with recent evidence suggesting critical involvement in semantic retrieval or selection (Gough, Nobre, & Devlin, 2005; Moss et al., 2005; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Wagner, Pare-Blagoev, Clark, & Poldrack, 2001). Broca's area has also been implicated in the recognition of hand and mouth actions, with recent evidence suggesting a key role in recognizing actions as part of the "mirror" or "observation–execution matching" system (Buccino et al., 2001; Buccino, Binkofski, & Riggio, 2004; Nishitani, Schurmann, Amunts, & Hari, 2005; Rizzolatti & Arbib, 1998; Sundara, Namasivayam, & Chen, 2001). The question we ask here is whether Broca's area processes speech-associated gestures as part of a language comprehension system (involving, in particular, semantic retrieval and selection), or as part of an action recognition system.

We begin by describing Broca's area in detail, including its various subdivisions and the functional roles attributed to these subdivisions. We then turn to the possible role or roles that Broca's area plays in processing speech-associated gestures.

2. Broca's area

2.1. Anatomy and connectivity of Broca's area

Broca's area in the left hemisphere and its homologue in the right hemisphere are designations usually used to refer to the pars triangularis (PTr) and pars opercularis (POp) of the inferior frontal gyrus. The PTr is immediately dorsal and posterior to the pars orbitalis and anterior to the POp. The POp is immediately posterior to the PTr and anterior to the precentral sulcus (see Fig. 1). The PTr and POp are defined by structural landmarks that only probabilistically (see Amunts et al., 1999) divide the inferior frontal gyrus into anterior and posterior cytoarchitectonic areas 45 and 44, respectively, by Brodmann's classification scheme (Brodmann, 1909). The anterior area 45 is granular, containing a layer IV, whereas the more posterior area 44 is dysgranular and distinguished from the more posterior agranular area 6 in that it does not contain Betz cells, i.e., in layer V.

These differences in cytoarchitecture between areas 45 and 44 suggest a corresponding difference in connectivity between the two areas and the rest of the brain. Indeed, area 45 receives more afferent connections from prefrontal cortex, the superior temporal gyrus, and the superior temporal sulcus, compared to area 44, which tends to receive more afferent connections from motor, somatosensory,
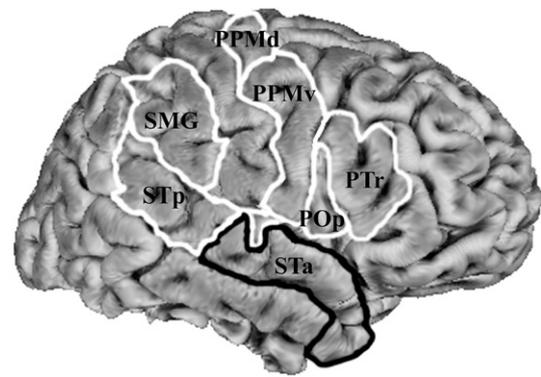

Fig. 1. Regions of interest used for structural equation modeling (SEM): POp, the pars opercularis of the inferior frontal gyrus; PTr, the pars triangularis of the inferior frontal gyrus; PPMv, ventral premotor and primary motor cortex; PPMd, dorsal premotor and primary motor cortex; SMG, the supramarginal gyrus of inferior parietal lobule; STp, superior temporal cortex posterior to primary auditory cortex and; STa, superior temporal cortex anterior to primary auditory cortex, extending to the temporal pole. Note: this schematic representation of regions is arbitrarily displayed on a right hemisphere and SEMs were performed on data averaged over hemispheres.

and inferior parietal regions (Deacon, 1992; Petrides & Pandya, 2002).

Taken together, the differences between areas 45 and 44 in cytoarchitecture and in connectivity suggest that these areas might perform different functions. Indeed, recent neuroimaging studies have been used to argue that the PTr and POp, considered here to probabilistically correspond to areas 45 and 44, respectively, play different functional roles in the human with respect to language comprehension and action recognition/understanding.

2.2. The role of Broca's area in language comprehension

The importance of Broca's area in language processing has been recognized since Broca reported impairments in his patient Leborgne (Broca, 1861). Indeed, for a long time, it was assumed that the role of Broca's area was more constrained to language production than language comprehension (e.g., Geschwind, 1965). The specialized role of Broca's area in controlling articulation *per se*, however, is questionable (Blank, Scott, Murphy, Warburton, & Wise, 2002; Dronkers, 1998; Dronkers, 1996; Knopman et al., 1983; Mohr et al., 1978; Wise, Greene, Buchel, & Scott, 1999). More recent evidence demonstrates that Broca's area is likely to play as significant a role in language comprehension as it does in language production (for review see Bates, Friederici, & Wulfeck, 1987; Poldrack et al., 1999; Vigneau et al., 2006).

More specifically, studies using neuroimaging and transcranial magnetic stimulation (TMS) of the PTr in both hemispheres yield results suggesting that this area plays a functional role in semantic processing during language comprehension. In particular, the PTr has been argued to play a role in controlled retrieval of semantic knowledge

(e.g., Gough et al., 2005; Wagner et al., 2001) or in selection among competing alternative semantic interpretations (e.g., Moss et al., 2005; Thompson-Schill et al., 1997).

If the PTr is involved in semantic retrieval or selection, then it should be highly active during instances of high lexical or sentential ambiguity. And it is—Rodd and colleagues (2005) recently found in two functional magnetic resonance imaging (fMRI) experiments that sentences high in semantic ambiguity result in more activity in the inferior frontal gyrus at a location whose center of mass is in the PTr.

In contrast to the functional properties of the PTr, the POp in both hemispheres has been argued to be involved in integrating or matching acoustic and/or visual information about mouth movements with motor plans for producing those movements (Gough et al., 2005; Hickok & Poeppel, 2004; Skipper, Nusbaum, & Small, 2005; Skipper, van Wassenhove, Nusbaum, & Small, 2007). For this and other reasons, the POp has been suggested to play a role in phonetic processing (see Skipper, Nusbaum, & Small, 2006 for a review). Specifically, the POp and other motor areas have been claimed (Skipper et al., 2005, 2006, 2007) to contribute to the improvement of phonetic recognition when mouth movements are observed during speech perception (see Grant & Greenberg, 2001; Reisberg, McLean, & Goldfield, 1987; Risberg & Lubker, 1978; Sumby & Pollack, 1954).

To summarize, there is reason to suspect that the divisions between the PTr and POp correspond to different functional roles in language processing. Specifically, the PTr becomes more active as semantic selection or retrieval demands are increased, whereas the POp becomes more active as demands for the integration of observed mouth movements into the process of speech perception increase.

2.3. The role of Broca's area in action recognition and production

In addition to these language functions, both the PTr and POp bilaterally have been proposed to play a functional role in the recognition, imitation, and production of actions (for review see Nishitani et al., 2005; Rizzolatti & Craighero, 2004). Although no clear consensus has been reached, there is some suggestion that these two brain areas play functionally different roles in action processing (Grezes, Armony, Rowe, & Passingham, 2003; Molnar-Szakacs et al., 2002; Nelissen, Luppino, Vanduffel, Rizzolatti, & Orban, 2005).

The functions of action recognition, imitation, and production are thought to have phylogenetic roots, in part because of homologies between macaque premotor area F5 and the POp of Broca's area (Rizzolatti, Fogassi, & Gallese, 2002). F5 in the macaque contains "mirror neurons" that discharge not only when performing complex goal-directed actions, but also when observing and imitating the same actions performed by another individual (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Kohler et al.,

2002; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Similar functional properties have been found in the human POp and Broca's area more generally, suggesting that Broca's area may be involved in a mirror or observation–execution matching system (Buccino et al., 2001; Buccino et al., 2004; Nishitani et al., 2005; Rizzolatti & Arbib, 1998; Rizzolatti & Craighero, 2004; Sundara et al., 2001).

A growing number of studies posit a link between Broca's area's involvement in language and its involvement in action processing (e.g., Floel, Ellger, Breitenstein, & Knecht, 2003; Hamzei et al., 2003; Iacoboni, 2005; Nishitani et al., 2005; Watkins & Paus, 2004; Watkins, Strafella, & Paus, 2003). Parsimony suggests that the anatomical association between a language processing area and a region involved in behavioral action recognition, imitation, and production ought to occur for a non-arbitrary reason. One hypothesis is that Broca's area plays a role in sequencing the complex motor acts that underlie linguistic *and* non-linguistic actions and, by extension, a role in understanding the sequence of those acts when performed by another person (Burton, Small, & Blumstein, 2000; Gelfand & Bookheimer, 2003; Nishitani et al., 2005).

Despite this overlap between the functional roles that Broca's area plays in language and action processing, most neuroimaging research on the human "mirror system" has focused on observable actions that are *not* communicative or are not typical of naturally occurring communicative settings. For example, critical tests of the mirror system hypothesis in humans have involved simple finger and hand movements (Buccino et al., 2001; Buccino et al., 2004; Iacoboni et al., 1999), object manipulation (Buccino et al., 2001; Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995), pantomime (Buccino et al., 2001; Fridman et al., 2006; Grezes et al., 2003), and lip reading and observation of face movements in isolation of spoken language (Buccino et al., 2001; Mottonen, Jarvelainen, Sams, & Hari, 2005; Nishitani & Hari, 2002).

2.4. The role of Broca's area in processing speech-associated gestures

Given that speech-associated gestures are relevant to language comprehension and are themselves observable actions, we can make two sets of predictions regarding the role of Broca's area in processing spoken language accompanied by gestures, displayed in Tables 1A and B, respectively.

Based on the hypothesized role that Broca's area, specifically the PTr, plays in semantic processing of words and sentences, we can derive the following set of predictions (see Table 1A). If the PTr is important in resolving ambiguity that arises in comprehending spoken words and sentences, reducing ambiguity should reduce the role of this area during speech comprehension. Given that gesture often provides a converging source of semantic information in spoken language (Goldin-Meadow, 2003; McNeill, 1992) that improves comprehension (see previously cited

Table 1
Predictions regarding the influence of Broca's area on other brain areas during the *Gesture*, *Self-Adaptor*, *No-Hand-Movement*, and *No-Visual-Input* conditions

Hypothesized roles of Broca's area	Condition			
	Gesture	*Self-Adaptor*	*No-Hand-Movement*	*No-Visual-Input*
(A) Language comprehension				
PTr: semantic selection or retrieval	+	++++	++++	++++
POp: using face movements as related to phonology	+	++++	++++	+
(B) Action recognition and production	++++	+++	++	+

The predictions are based on whether Broca's area processes speech-related gestures as part of a language comprehension system (involving, in particular, semantic retrieval or selection, (A), or as part of an action recognition and production system (B). "PTr" refers to the pars triangularis and "POp" refers to the pars opercularis. The plus signs indicate the relative strength of the relationship between Broca's area and other areas, with "+" signifying a weak relationship and "++++" a strong relationship.

references), the presence of gesture should reduce the ambiguity of speech (see Holler & Beattie, 2003). Thus, when speech-associated gestures are present, compared to when they are not, the PTr should have reduced influence on other brain areas ("+" in the PTr row in Table 1A for the *Gesture* condition).

To the extent that message-level information is clarified by the presence of gesture, there should be a reduced need to attend closely to the phonological content of speech since attention is focused on the meaning of the message, rather than its phonological form. Thus, if speech-associated gestures reduce message ambiguity, then there should also be a reduced influence of the POp on language comprehension areas because there will be less need to integrate acoustic and visual information with motor plans in the service of phonology ("+" in the POp in Table 1A for the *Gesture* condition).

In contrast, when the face is moving (i.e., talking) and the hands are moving in a way that is not meaningful in relation to the spoken message, or the face is moving and the hands are not moving, the POp should be more involved with other cortical areas because it relies on face movements to help decode phonology from speech ("++++" in the POp row in Table 1A for the *Self-Adaptor* and *No-Hand-Movement* conditions). By similar reasoning, when the face is moving and the hands are moving in ways that are not meaningful with respect to speech, the face is moving and the hands are not moving, or when there is no visual input, the PTr should interact with other brain areas more strongly ("++++" in the PTr row in Table 1A for the *Self-Adaptor, No-Hand-Movement*, and *No-Visual-Input* conditions). In other words, when there are no speech-associated gestures, there is less converging semantic information to aid in comprehension of the spoken message. As a result, there will be a greater need for the PTr to aid in the interpretation process through semantic retrieval or selection.

To summarize the first set of predictions, as outlined in Table 1A, if Broca's area processes speech-associated gestures in accord with its role in language comprehension (i.e., the PTr in retrieval/selection and the POp in phonol-

ogy), then Broca's area should have *less* influence on other brain areas when processing stories accompanied by speech-associated gestures (one plus) than when processing stories accompanied by self-grooming movements or by no hand movements or by no visual input at all (many plusses).

However, if Broca's area is processing speech-associated gesture as part of an action recognition system, we arrive at a different set of predictions (see Table 1B). From this perspective, Broca's area should show a greater influence on other brain areas when speech-associated gestures are present, compared to when they are not. The rationale here is that speech-associated gestures are important goal-directed actions that aid in the goal of communication, namely comprehension. Thus, it would be expected that the PTr and POp should both have a greater influence on other brain areas in the presence of speech-associated gestures because there is an increased demand on the mirror or observation–execution matching functions of the human mirror system ("++++" in Table 1B for the *Gesture* condition). By similar reasoning, Broca's area should have increasingly less influence on other brain regions when accompanied by face movements and hand movements that are non-meaningful with respect to the speech, face movements alone, or no visual input at all ("+++", "++", and "+" in Table 1B for the *Self-Adaptor, No-Hand-Movement*, and *No-Visual-Input* conditions, respectively). That is, as the number of goal-directed actions decreases, there should be a concomitant decrease in mirror or observation–execution matching functions of the human mirror system because there are fewer movements to mirror or match.

To summarize the second set of predictions, as outlined in Table 1B, if Broca's area processes speech-associated gestures in accord with its role in action recognition (i.e., as part of a mirror or observation–execution matching system), it should have *more* influence on other brain areas when processing stories accompanied by speech-associated gestures (many plusses) than when processing stories accompanied by self-grooming or no hand movements or no visual input at all (fewer plusses).

2.5. The mirror system and speech-associated gestures

We hypothesized that the influence of Broca's area on the rest of the brain when speech-associated gestures are observed will be more consistent with the first set of predictions than the second, i.e., with a role primarily in the service of semantic retrieval or selection and phonology than of action recognition (see Table 1). If this hypothesis is correct, then the question becomes—which regions are serving the action recognition functions posited by the mirror or observation–execution matching account of Broca's area? This subsection addresses this question, and our predictions are displayed in Table 2.

The hypothesis that the language comprehension account better explains the influence of Broca's area's on the rest of the brain when speech-associated gestures are observed grew out of previous neuroimaging work in our laboratory on the neural systems involved in listening to spoken language accompanied only by face movements. In this research, we showed that when a speaker's mouth is visible, the motor and somatosensory systems related to production of speech are more active than when it is not visible. In particular, the ventral premotor and primary motor cortices involved in making mouth and tongue movements (PPMv; see Fig. 1) and the posterior superior temporal cortices (STp; see Fig. 1) show particular sensitivity to visual aspects of observed mouth movements (Skipper et al., 2005, 2007).

By analogy to this previous work, we predict that the PPMv and dorsal premotor and primary motor cortex (PPMd; see Fig. 1), *both* involved in producing hand and arm movements (e.g., Schubotz & von Cramon, 2003), will be sensitive to observed speech-associated gestures (see Table 2). In our previous research, interactions between PPMv and STp (which is involved in phonological aspects of speech perception and production, Buchsbaum, Hickok, & Humphries, 2001) were associated with perception of speech sounds, presumably because some face movements are correlated with phonological aspects of speech perception. Again, by analogy, activity in the PPMv and PPMd should influence other brain areas involved in generating hand movements, such as the supramarginal gyrus (SMG; see Fig. 1) of the inferior parietal lobule (Harring-

ton et al., 2000; Rizzolatti, Luppino, & Matelli, 1998). However, activity in the PPMv and PPMd should also influence areas involved in understanding the meaning of language because speech-associated gestures are correlated with semantic aspects of spoken language comprehension. Recent research has implicated the superior temporal cortex anterior to Heschel's Gyrus (STa; see Fig. 1) in comprehension of spoken words, sentences, and discourse (see Crinion & Price, 2005; Humphries, Love, Swinney, & Hickok, 2005; Humphries, Willard, Buchsbaum, & Hickok, 2001; Vigneau et al., 2006) and, specifically, the interaction between grammatical and semantic aspects of language comprehension (Vandenberghe, Nobre, & Price, 2002).

To summarize, as shown in Table 2 (left column), we hypothesize that interactions among the PPMv, PPMd, SMG, and STa may underlie the effects of speech-associated gestures on the neural systems involved in spoken language comprehension. We argue by analogy with our previous research that it is interaction among these areas, rather than processing associated with Broca's area *per se*, that constitutes the mirror or observation–execution matching system associated with processing speech-associated gestures. If it is gesture's semantic properties (rather than its properties as a goal-directed hand movement) that are relevant to Broca's area, then we should expect Broca's area to have relatively little influence on other cortical areas when listeners are given stories accompanied by speech-associated gestures (see above and Table 1A). That is, if speech-associated gestures reduce the need for semantic selection/retrieval and the need to make use of face movements in service of phonology, then Broca's area should have relatively little influence on the PPMv, PPMd, SMG, and STa when gestures are processed (as opposed to other hand movements that are not meaningful with respect to spoken content).

In contrast, as shown in Table 2 (right column), we hypothesize that interactions among the POp, PPMv, and STp underlie the effects of face movements on the neural systems involved in spoken language comprehension. That is, based on previous research, interaction among these areas, including Broca's area to the extent that Broca's area (i.e., the POp) plays a role in phonology, constitutes the

Table 2

Hypothesized regions of interest (ROIs; see Fig. 1 caption for the definition of ROI abbreviations and Fig. 1 for location of ROIs) that constitute the "mirror system" associated with processing observed speech-associated gestures or face movements

ROIs	"Mirror system" associated with	
	Speech-associated gestures (*Gesture*)	Face movements (*Self-Adaptor*, *No-Hand-Movement*, and *No-Visual-Input*)
POp		*
PPMv	*	*
STp		*
PPMd	*	
SMG	*	
STa	*	

An asterisk indicates that a region is predicted to be part of a "mirror system" involved in processing a movement. PPMv appears in both networks because it plays a role in both hand and mouth movements. See Section 2 for further explanation.

mirror or observation–execution matching system associated with processing face movements. Thus, these areas should constitute the mirror or observation–execution matching system when speech-associated gestures are not observed, i.e., when the face is moving and the hands are moving in a way that is not meaningful in relation to the spoken message, or the face is moving and the hands are not moving.

To test the predictions outlined in Tables 1 and 2, we performed fMRI while participants listened to adapted Aesop's Fables without visual input (*No-Visual-Input* condition) or during three conditions with a video of the storyteller whose face and arms were visible. In the *No-Hand-Movement* condition, the storyteller kept her arms in her lap so that the only visible movements were her face and mouth. In the *Gesture* condition, she produced speech-associated gestures that bore a relation to the semantic content of the speech they accompanied (these were metaphoric, iconic, and deictic gestures, McNeill, 1992). In the *Self-Adaptor* condition, she produced self-grooming movements that were not meaningful with respect to the story (e.g., scratching herself or adjusting her clothes, hair, or glasses). As our hypotheses are inherently specified in terms of relationships among brain regions, we used structural equation models (SEMs) to analyze the strength of association of patterns of activity in the brain regions enumerated above (i.e., PTr, POp, PPMv, PPMd, SMG, STp, and STa) in relation to these four conditions (i.e., *Gesture*, *Self-Adaptor*, *No-Hand-Movement*, and *No-Visual-Input*).

3. Methods

3.1. Participants

Participants were 12 (age 21 ± 5 years; 6 females) right-handed (as determined by the Edinburgh handedness inventory; Oldfield, 1971) native English speakers who had no early exposure to a second language. All participants had normal hearing and vision with no history of neurological or psychiatric illness. Participants gave written informed consent and the Institutional Review Board of the Biological Science Division of The University of Chicago approved the study.

3.2. Stimuli and task

As described above, participants listened to adapted Aesop's Fables in *Gesture, Self-Adaptor*, *No-Hand-Movement*, and *No-Visual-Input* conditions. The storyteller rehearsed her performance before telling the story with gestures, self-adaptors, and no hand movements. We used rehearsed stimuli to keep the actress's speech productions constant across the four conditions. Hand and arm movements during speech production (or their lack) can change dimensions of the speech signal, such as prosody, lexical content, or timing of lexical items. The actress practiced

the stimuli so that her prosody was consistent across stimuli, and lexical items were the same and occurred in the same temporal location across stimuli. The *No-Visual-Input* stimuli were created by removing the video track from the *Gesture* condition; thus the speech in these two conditions was identical.

Another reason that we used rehearsed stimuli was to be sure that the self-adaptor movements occurred in the same temporal location as the speech-associated gestures. The speech-associated gestures themselves were modeled after natural retellings of the Aesop's Fables. The self-adaptor movements were rehearsed so that they occurred in the same points in the stories as the speech-associated gestures. Thus, the *Gesture* and *Self-Adaptor* conditions were matched for overall movement.

Each story lasted 40–50 s, and participants were asked to listen attentively. Each condition was presented once in a randomized manner in each run. There were two runs lasting approximately 4 min each. Participants heard a total of eight stories, two in each condition, and did not hear the same story more than once. Conditions were separated by a baseline period of 12–14 s. During baseline and the *No-Visual-Input* condition, participants saw only a fixation cross, but they were not explicitly asked to fixate. Stories were matched and counter-balanced so that the *Gesture* condition could be compared to the *Self-Adaptor*, *No-Hand-Movement*, and *No-Visual-Input* conditions. For example, one group of participants heard story 1 in the *Gesture* condition and story 2 in the *Self-Adaptor* condition; the matched group heard story 1 in the *Self-Adaptor* condition and story 2 in the *Gesture* condition.

Audio was delivered at a sound pressure level of 85 dB-SPL through headphones containing MRI-compatible electromechanical transducers (Resonance Technologies, Inc., Northridge, CA). Participants viewed video stimuli through a mirror attached to the head coil that allowed them to see a screen at the end of the scanning bed. Participants were monitored with a video camera.

Following the experiment, participants were asked true and false questions about each story to assess (1) whether they paid attention during scanning, and (2) whether they could answer content questions when listening to stories in the *Gesture* condition more accurately than when listening to stories in the other conditions.

3.3. Imaging parameters

Functional imaging was performed at 3 T (TR = 2 s; TE = 25 ms; FA = 77°; 30 sagittal slices; $5 \times 3.75 \times 3.75$ mm voxels) with BOLD fMRI (GE Medical Systems, Milwaukee, WI) using spiral acquisition (Noll, Cohen, Meyer, & Schneider, 1995). A volumetric T1-weighted inversion recovery spoiled grass sequence was used to acquire images on which anatomical landmarks could be found and functional activation maps could be superimposed.

3.4. Data analysis

Functional images were spatially registered in three-dimensional space by Fourier transformation of each of the time points and corrected for head movement, using the AFNI software package (Cox, 1996; http://afni.nimh.nih.gov/afni/). Scanner-induced spikes were removed from the resulting time series, and the time series was linearly and quadratically detrended. Time series data were analyzed using multiple linear regression. There were separate regressors of interest for each of the four conditions (i.e., *Gesture*, *Self-Adaptor*, *No-Hand-Movement*, and *No-Visual-Input*). These regressors were waveforms with similarity to the hemodynamic response, generated by convolving a gamma-variant function with the onset time and duration of the blocks of interest. The model also included a regressor for the mean signal and six motion parameters, obtained from the spatial alignment procedure, for each of the two runs. The resulting *t*-statistics associated with each condition were corrected for multiple comparisons to $p < .05$ using a Monte Carlo simulation to optimize the relationship between the single voxel statistical threshold and the minimally acceptable cluster size (Forman et al., 1995). The time series was mean corrected by the mean signal from the regression.

Next, cortical surfaces were inflated (Fischl, Sereno, & Dale, 1999) and registered to a template of average curvature (Fischl, Sereno, Tootell, & Dale, 1999) using the Freesurfer software package (http://surfer.nmr.mgh.harvard.edu). The surface representations of each hemisphere of each participant were then automatically parcellated into regions of interest (ROIs) that were manually subdivided into further ROIs (Fischl et al., 2004). There were seven ROIs per participant in the final analysis (Fig. 1). These regions were chosen because they (1) operationally comprise Broca's area (i.e., the POp and PTr), (2) we have previously shown them to be involved in producing mouth movements and to underlie the influence of observable mouth movements on speech perception (i.e., the PPMv and STp), and (3) they were hypothesized to be involved in producing speech-associated gestures or to underlie the influence of observed speech-associated gestures on comprehension (i.e., the PPMv, PPMd, SMG, and STa; see Section 2 for further details and references).

The POp was delineated anteriorly by the anterior ascending ramus of the sylvian fissure, posteriorly by precentral sulcus, ventrally by the Sylvian fissure, and dorsally by the inferior frontal sulcus. The PTr was delineated anteriorly by the rostral end of the anterior horizontal ramus of the Sylvian fissure, posteriorly by the anterior ascending ramus of the Sylvian fissure, ventrally by the anterior horizontal ramus of the Sylvian fissure, and dorsally by the inferior frontal sulcus.

The PPMv was delineated anteriorly by the precentral sulcus, posteriorly by the anterior division of the central sulcus into two halves, ventrally by the posterior horizontal ramus of the Sylvian fissure to the border with insula cor-

tex, and dorsally by a line extending the superior aspect of the inferior frontal sulcus through the precentral sulcus, gyrus, and central sulcus. The PPMd was delineated anteriorly by the precentral sulcus, posteriorly by the anterior division of the central sulcus into two even halves, ventrally by a line extending the superior aspect of the inferior frontal sulcus through the precentral sulcus, gyrus, and central sulcus, and dorsally the most superior point of the precentral sulcus. Both premotor and primary motor cortex were included in PPMv and PPMd because the somatotopy in premotor and primary motor cortex is roughly parallel (e.g., Godschalk, Mitz, van Duin, & van der Burg, 1995). The use of the inferior frontal sulcus to determine the boundary between the PPMv and PPMd derives from previous work in our lab (Hluštík, Solodkin, Skipper, & Small, in preparation) showing that multiple somatotopic maps exist in the human which are roughly divisible into a ventral section containing face and hand representations and a dorsal section containing hand and arm and leg representations (see also Fox et al., 2001; Schubotz & von Cramon, 2003).

The SMG was delineated anteriorly by the postcentral sulcus, posteriorly by the angular gyrus, ventrally by posterior horizontal ramus of the Sylvian fissure, and dorsally by the intraparietal sulcus. The STp was delineated anteriorly by Heschel's sulcus, posteriorly by a coronal plane defined as the endpoint of the Sylvian fissure, ventrally by the upper bank of the superior temporal sulcus, and dorsally by the posterior horizontal ramus of the Sylvian fissure. Finally, the STa was delineated anteriorly by the temporal pole, posteriorly by Heschel's sulcus, ventrally by the dorsal aspect of the upper bank of the superior temporal sulcus, dorsally by a posterior horizontal ramus of the Sylvian fissure.

Following parcellation into ROIs, the coefficients, corrected *t*-statistic associated with each regression coefficient and contrast, and time series data were interpolated from the volume domain to the surface representation of each participant's anatomical volume using the SUMA software package (http://afni.nimh.nih.gov/afni/suma/). A relational database was created in MySQL (http://www.mysql.com/) and individual tables were created in this database for each hemisphere of each participant's coefficients, corrected *t*-statistics, time series, and ROI data.

The *R* statistical package was then used to analyze the information stored in these tables (Ihaka & Gentleman, 1996; http://www.r-project.org/). First, *R* was used to conduct a group-based ANOVA to determine baseline levels of activation for each condition. This ANOVA took as input the coefficients from each individual's regression model and had one fixed factor, Condition, and one random factor, Participant. Condition had four levels, one for each condition. Following this ANOVA, post hoc contrasts between conditions were performed.

Next, we used *R* to query the database to extract from each of the seven ROIs the time series for each condition of only those surface nodes that were active in at least

one of the four conditions for each hemisphere of each participant. A node was determined to be active if any of the *Gesture*, *Self-Adaptor*, *No-Hand-Movement*, or *No-Visual-Input* conditions was active at $p < .05$, corrected. In each of the four resulting time series, time points with signal change values greater than 10% were replaced with the median signal change. The resulting time series corresponding to each of the active nodes for each condition in each of the ROIs for each hemisphere of each participant was averaged. Finally, the resulting four time series were averaged over participants and hemisphere, thus establishing for each ROI one representative time series for each of the four conditions.

After this, the second derivative of the time series was calculated for each condition. The second derivative of the time series was used for further analysis because the second derivative detects peaks in the time series that reflect events in the stories (Skipper, Goldin-Meadow, Nusbaum, & Small, in preparation). The general "boxcar" shape of the time series associated with stories obscures these fluctuations, as they tend to be very similar across conditions.

3.5. Structural equation modeling

Structural equation models (SEM) are multivariate regression models that are being used to study systems level neuroscience (for a review see Buchel & Friston, 1997; Horwitz, Tagamets, & McIntosh, 1999; McIntosh & Gonzalez-Lima, 1991, 1992, 1993; Penny, Stephan, Mechelli, & Friston, 2004; Rowe, Friston, Frackowiak, & Passingham, 2002; Solodkin, Hlustik, Chen, & Small, 2004). SEMs comprise two components, a measurement model and a structural model. The measurement model relates observed responses to latent variables and sometimes to observed covariates. The structural model then specifies relations among latent variables and regressions of latent variables on observed variables. Parameters are estimated by minimizing the difference between the observed covariances in the measurement model and those implied by the structural model.

In SEMs applied to neuroscience data, the measurement model is based on observed covariances of time series data from ROIs (described above) and the structural model is inferred from the known connectivity of primate brains. SEM equations are solved simultaneously using an iterative maximum likelihood method. The best solution to the set of equations minimizes the differences between the observed covariance from the measurement model and the predicted covariance matrices from the structural model. A χ^2 measure of goodness of fit between the predicted and observed covariance matrices is determined. If the null hypothesis is *not* rejected (i.e., $p < .05$), a good fit was obtained. The result is a connection weight (or path weight) between two connected regions that represents the influence of one region on the other, controlling for the influences of the other regions in the structural model.

3.6. Model uncertainty and Bayesian model averaging of structural equation models

SEMs are typically proposed and tested without consideration that model selection is occurring. That is, when a reasonable χ^2 is found based on a theoretical model, rarely is it acknowledged that alternative models exist that could also yield a reasonable χ^2 with, perhaps, substantially different connection weights. This is a form of model uncertainty. An alternative approach that accounts for model uncertainty is Bayesian model averaging (Hoeting, Madigan, Raftery, & Volinsky, 1999; Kass & Raftery, 1995). This procedure involves averaging over *all* competing models, thus producing more reliable and stable results and providing better predictive ability than using any single model (Madigan & Raftery, 1994).

Testing all possible models, however, is not feasible in the typical laboratory setting because it requires prohibitive computation time, as the number of possible models increases more than exponentially with the number of nodes in the model (Hanson et al., 2004). Hanson et al. (2004) estimate that exhaustive search for all possible SEMs comprised of eight brain areas could require up to 43 centuries.

Thus, instead of attempting to perform all possible SEMs on all of our ROIs, we chose to perform exhaustive search on three smaller models consisting of five ROIs selected from the seven ROIs for the *Gesture*, *Self-Adaptor*, *No-Hand-Movement*, and *No-Visual-Input* conditions. These models consider the STp, SMG, and STa as "hubs" to look at the connections weights of these hubs with Broca's area (i.e., the POp and PTr) and premotor and primary cortex (i.e., the PPMv and PPMd). We considered the STp, SMG, and STa to be hubs because each of these areas has been associated with speech perception and language comprehension in the past. That is, as described in Section 2, our predictions are about the influence that Broca's area and motor areas have on areas involved in speech perception and language comprehension when stories with gestures, self-adaptors, no hand movements and no visual input are processed. We, therefore, kept Broca's area and motor areas constant in the models and varied the speech perception and language comprehension hubs to see the impact of Broca's area and motor areas on speech perception and language comprehension regions during the various conditions. Thus, we produced 3 sets of Bayesian averaged SEMs. Each set contained up to eight physiological plausible connections between (1) STp and POp, PTr, PPMv, or PPMd, (2) SMG and POp, PTr, PPMv, or PPMd, and (3) STa and POp, PTr, PPMv, or PPMd (Deacon, 1992; Petrides & Pandya, 2002).

Our SEMs were solved using the SEM package written by J. Fox for *R* (http://cran.r-project.org/doc/packages/sem.pdf). Forward and backward connections between two regions were solved independently and not

simultaneously, though every possible model with connections in each direction was tested. Thus, a total of 38,002 models for each hub were tested, resulting in a total 114, 006 tested models. All models were solved using up to 150 processors on a grid-computing environment (Hasson, Skipper, Wilde, Nusbaum, & Small, submitted for publication). For each of the three models, the SEM package was provided the correlation matrix derived from the second derivative of the time series (see above) between all regions within that model for each of the four conditions. Only models whose χ^2 was not significant (i.e., models demonstrating a good fit; $p > .05$) were saved.

The resulting path weights from each of the saved well-fitting models from each of the three models were averaged using Bayesian model averaging, resulting in one model for each hub (i.e., independent models for the STp, SMG, or STa). Bayesian model averaging consists of averaging weighted by the Bayesian information criterion for each model (BIC; Hoeting et al., 1999). The Bayesian information criterion adjusts the χ^2 for the number of parameters in the model, the number of observed variables, and the sample size.

Specifically, averaging of path weights was performed according to the following formulas (adapted from Hoeting et al., 1999):

$$E(\alpha|y) = \sum_{k=1}^{n} P(M_k|y) E(\beta|y, M_k)$$

$$\mathrm{Var}(\alpha|y) = \sum_{k=1}^{n} P(M_k|y)(\mathrm{Var}(\beta|y, M_k) + E^2(\beta|y, M_k)) - E^2(\alpha|y)$$

where:

$P(M_k|y)$ is the posterior probability of the model M_k given the data y, and

$E(\beta|y, M_k)$ is the model-specific estimate of β, the path weight.

The posterior probability $P(M_k|y)$ in the above formula for each model is estimated as

$$P(M_k|y) = \frac{P(M_k)\mathrm{e}^{-0.5\left(\mathrm{BIC}(M_k) - \overline{\mathrm{BIC}(M_k)}\right)}}{\sum_{k=1}^{n} P(M_k)\mathrm{e}^{-0.5\left(\mathrm{BIC}(M_k) - \overline{\mathrm{BIC}(M_k)}\right)}}$$

The prior probability $P(M_k)$ was assumed to be from the uniform distribution,

$$P(M_k) = \frac{1}{n}$$

Resulting connection weights between regions were compared between the different conditions within each of the three models independently using *t*-tests correcting for heterogeneity of variance and unequal sample sizes by the Games-Howell method (Kirk, 1995). Degrees of freedom for unequal sample sizes were calculated using Welch's method (Kirk, 1995; though the number of models for each condition were not significantly different).

4. Results

4.1. Behavioral

Mean accuracy for the true/false questions asked about the stories after scanning was 100%, 94%, 88%, and 84% for the *Gesture*, *Self-Adaptor*, *No-Hand-Movement*, and *No-Visual-Input* conditions, respectively. Participants were significantly more accurate at answering questions after hearing the *Gesture* story than after hearing the stories in the other three conditions ($t = 3.7$; $df = 11$; $p < .003$). Participants were significantly more accurate in answering questions after hearing stories in the *Gesture* condition, compared to the stories in the *No-Hand-Movement* ($t = 2.7$; $df = 11$; $p < .02$) and *No-Visual-Input* ($t = 2.4$; $df = 11$; $p < .03$) conditions. The difference in accuracy between the *Gesture* and *Self-Adaptor* conditions was not significant ($t = 1.5$; $df = 11$; $p < .15$).

4.2. Activation data

We have presented baseline contrasts of all conditions and contrasts between the *Gesture* condition and the other conditions elsewhere (Josse et al., in preparation, 2005; Josse, Skipper, Chen, Goldin-Meadow, & Small, 2006). In these analyses, we found that the inferior frontal gyrus, premotor cortex, superior temporal cortex, and inferior parietal lobule were active above a resting baseline in all conditions.

4.3. Structural equation models

The analyses presented here, represented by Figs. 2, 4, and 6, focus on the strongest and weakest connection weights for each of the three models. In each of these figures, the arrowed line indicates the connection strength and the direction of influence between an area and the area(s) to which it connects. Dotted arrowed lines indicate areas that have a negative influence on the area(s) to which they connect. Thick orange arrowed lines indicate connection weights that are statistically *stronger* for the condition in which that connection appears, compared to the same connection in all of the other conditions ($p < .00001$ in all cases). Similarly, thick blue arrowed lines indicate connection weights that are statistically *weaker* for the condition in which that connection appears, compared to the same connection in all of the other conditions ($p < .00001$ in all cases). Thin gray arrowed lines indicate connection weights that were not statistically different from the condition in which that connection appears, compared to the same connection in at least one of the other conditions.

Fig. 2 shows the result of the Bayesian averaging of connection weights for all models with the STp ROI connected with the POp, PTr, PPMv, and PPMd ROIs. With the exception of the connections between the POp and STp, the *Gesture* condition produced the
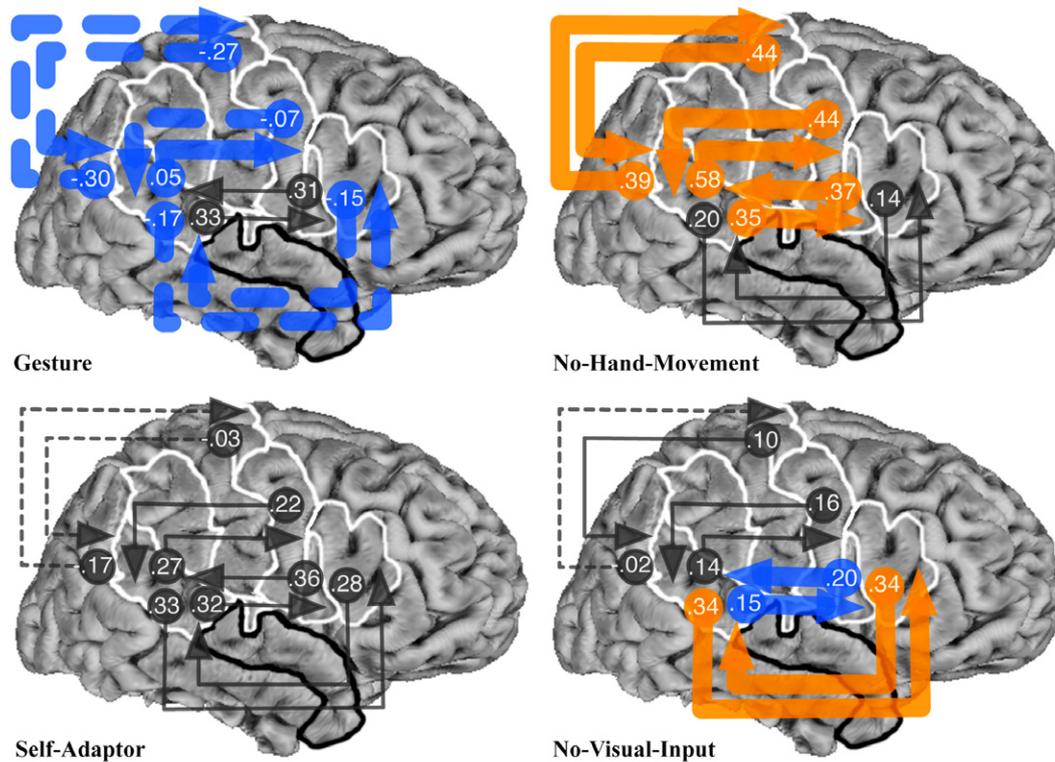
Fig. 2. Results of Bayesian averaging of connection weights from exhaustive search of all structural equation models with connections between STp and POp, PTr, PPMv, or PPMd regions of interest (ROIs), for the *Gesture*, *No-Hand-Movement*, *Self-Adaptor*, and *No-Visual-Input* conditions. See Fig. 1 caption for the definition of ROI abbreviations and Fig. 1 for location of ROIs. Arrowed lines indicate connections and the direction of influence of an area on the area(s) to which it connects. Connection weights are given at the beginning of each arrowed line. Dotted arrowed lines show areas that have a negative influence on the area(s) to which it connects. Thick orange arrowed lines indicate connection weights that are statistically stronger for the condition in which that connection appears, compared to the same connection in all of the other conditions ($p < .00001$ in all cases). Thick blue arrowed lines indicate connection weights that are statistically weaker for the condition in which that connection appears, compared to the same connection in all of the other conditions ($p < .00001$ in all cases). Thin gray arrowed lines indicate connection weights that were not statistically different from the condition in which that connection appears, compared to the same connection in at least one of the other conditions.

statistically weakest connection weights for all STp connections (note the blue arrows). This is consistent with our prediction that Broca's area should have a reduced role when speech-associated gestures are processed. When speech-associated gestures are present, listeners focus more on the meaning of the message, jointly specified by speech and gesture, rather than on phonological information, as reflected in the functional neural associations.

In contrast, the *No-Hand-Movement* condition produced the strongest weights between all connections with the STp, with the exception of the connections between the STp and PTr (note the orange arrows). This result is consistent with our earlier research described above, showing that most of these regions are involved in recognizing observed mouth movements in the service of speech perception. Without speech-associated gestures, the only source of converging information about the speech in the *No-Hand-Movement* condition is the talker's mouth movements.

The *Self-Adaptor* condition produced no significantly greater or weaker connection weights among any regions (note the gray arrows). Finally, the *No-Visual-Input* condi-

tion produced the strongest connection weights between the STp and PTr, and the weakest connection weights between the STp and POp. Again, this is consistent with our initial prediction: face movements were not observable and, thus, the influence of the POp should be reduced. In contrast, the influence of the PTr may be greater because of its increased role in controlled semantic interpretation, given the absence of gesture.

To summarize, the results show that the *Gesture* condition produced on average the weakest connection weights between the STp and Broca's area, and between the STp and premotor and primary motor cortex (Fig. 3), consistent with our predictions (see Table 1A and Table 2). The *No-Visual-Input* condition resulted in the strongest connection weights between the STp and premotor and primary motor cortices.

Fig. 4 shows the result of Bayesian averaging of connection weights for all models with the SMG ROI connected with the POp, PTr, PPMv, and PPMd ROIs. The *Gesture* condition produced the statistically weakest connections between the SMG for three of the four connections with Broca's area. In contrast, the *Gesture* condition produced the strongest connection weights
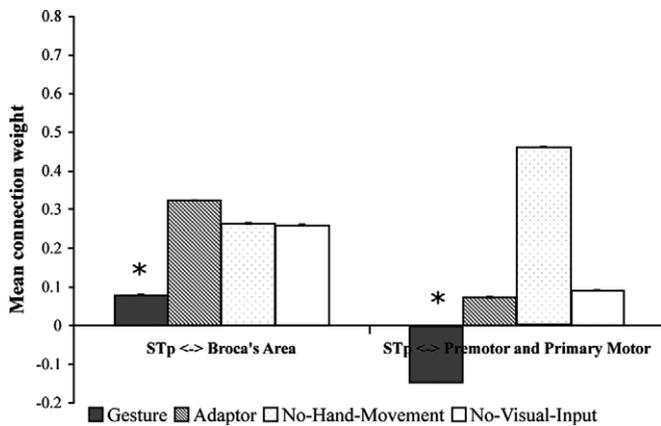
Fig. 3. Mean of all posterior superior temporal (STp) connection weights from Bayesian averaging from exhaustive search of all structural equation models shown in Fig. 2 with the pars opercularis and pars triangularis (i.e., Broca's area) and dorsal and ventral premotor and primary motor cortex. Asterisks indicate a significant difference between the *Gesture* condition and the *No-Hand-Movement*, *Self-Adaptor*, and *No-Visual-Input* conditions ($p < .00001$ in all cases). Error bars indicate standard error.

posed role of the SMG, PPMv, and PPMd in mirroring hand and arm movements. The *Self-Adaptor*, *No-Hand-Movement*, and *No-Visual-Input* conditions produced a strong negative influence, a weak influence, or occasionally a moderate influence between the SMG, PPMv, and PPMd areas.

On the other hand, the *Self-Adaptor* condition produced the strongest connections between the SMG and Broca's area, compared to the *Gesture*, *No-Hand-Movement*, and *No-Visual-Input* conditions. Although this finding appears inconsistent with the role that Broca's area was predicted to play in processing self-adaptors (see Table 1), it need not be. Self-adaptor movements are hand movements that *could* have provided information about semantic content if they have been seen as gestures rather than as self-grooming movements. Unlike spoken words, gesture production is not bound by the conventions of language and thus reflects a momentary cognitive construction, rather than the constraints of a culturally uniform linguistic system. Any particular hand movement that accompanies speech therefore has the potential to be a gesture, at least until it is recognized as a specific kind of speech-irrelevant action. Because listeners cannot know *a priori* that a self-adaptor does not provide message-relevant information, self-adaptors may place some demand on the listener's semantic selection or retrieval.

between the SMG and three of the four connections with premotor and primary motor areas (i.e., the PPMd and PPMv). This is again consistent with our prediction that Broca's area should play a reduced role when speech-associated gestures are observable and the pro-
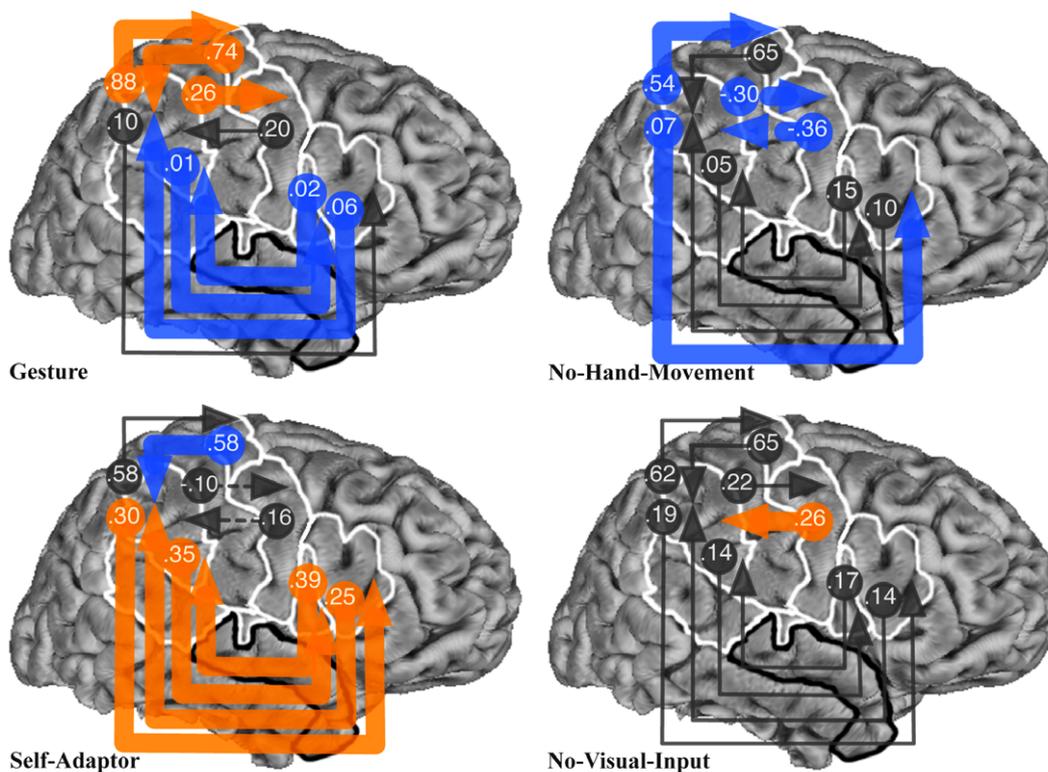

Fig. 4. Results of Bayesian averaging of connection weights from exhaustive search of all structural equation models with connections between SMG and POp, PTr, PPMv, or PPMd regions of interest (ROIs), for the *Gesture*, *No-Hand-Movement*, *Self-Adaptor*, and *No-Visual-Input* conditions. See Fig. 1 caption for the definition of ROI abbreviations and Fig. 1 for location of ROIs. See Fig. 2 caption for the meaning of arrowed lines and their colors.

In summary, results show that the *Gesture* condition produced the weakest average connection weights between the SMG and Broca's area, and the strongest connection weights between the SMG and premotor and primary motor cortex (Fig. 5), consistent with our predictions (see Table 1A and Table 2).

Fig. 6 shows the results of Bayesian averaging of connection weights for all models with the STa ROI connected with the POp, PTr, PPMv, and PPMd ROIs. The *Gesture* condition produced the weakest connections involving all STa and all Broca's area connections, again, consistent with initial predictions. On the other hand, the *Gesture* condition produced the strongest influence of the STa on the PPMv and of the PPMd on the STa. The *Self-Adaptor* condition produced the strongest connection weights between the STa and the PTr and from the STa to the PPMd, and the weakest connection weights from the PPMv to the STa. The *No-Hand-Movement* condition produced the strongest weights between the STa and the POp and from the PPMv to the STa. The *No-Visual-Input* condition produced the weakest connections between the STa and all but one of the premotor and primary motor cortex areas.

These results can be summarized as showing that the *Gesture* condition produced the weakest average connection weights between the STa and Broca's area, and the strongest connection weights between the STa and premotor and primary motor cortex (Fig. 7), consistent with our predictions (see Table 1A and Table 2).

Finally, to graphically summarize the results over all three sets of SEMs with respect to Broca's area, we averaged all connections with the POp and PTr (Fig. 8). This representation shows that the statistically weakest connection weights associated with the POp correspond to the *Gesture* condition (though not significantly different from

the *No-Visual-Input* condition; Fig. 8a). Similarly, the statistically weakest connection weights associated with the PTr correspond to the *Gesture* condition (Fig. 8b). That is, Broca's area produces the weakest connection weights in the Gesture condition, consistent with the predictions summarized in Table 1A.

5. Discussion

We have shown that Broca's area, defined as the POp and PTr, generally has the weakest impact on other motor and language-relevant cortical areas (i.e., PPMv, PPMd, SMG, STp, and STa) when speech is interpreted in the context of meaningful gestures, as opposed to in the context of self-grooming movements, resting hands, or no visual input at all. These differences in connection strengths cannot simply be attributed to varying amounts of visually compelling information in the different conditions. The *Gesture* and *Self-Adaptor* conditions were matched on amount of visual motion information, and hand and arm movements occurred in approximately the same points in the stories in the two conditions. Thus, the differences are more likely to be a function of the *type of information* carried by the observed movements. Similarly, these differences cannot easily be attributed to varying levels of attention across conditions. Most accounts of neural processing associated with attention show *increased* levels of activity as attention or processing demands increase (e.g., Just, Carpenter, Keller, Eddy, & Thulborn, 1996). Yet, we show that the *Gesture* condition results in the *weakest* connection weights between the PTr and POp and other cortical areas; these low levels are concomitant with an increase in behavioral accuracy for the comprehension questions pertaining to the *Gesture* condition.

5.1. Speech-associated gestures and the role of Broca's area in language comprehension and action recognition

Why should the functional connectivity of Broca's area with other cortical areas be modulated by the presence or absence of speech-associated gestures? As reviewed in Section 2, our results were predicted by one particular account of the role of Broca's area in language comprehension—the view that Broca's area is important for semantic processing. The PTr has been hypothesized to be involved in the retrieval or selection of semantic information, and the POp has been hypothesized to be involved in matching acoustic and/or visual information about mouth movements with motor plans for producing those movements (Gough et al., 2005; Poldrack et al., 2001; Poldrack et al., 1999; Skipper et al., 2006; Thompson-Schill, Aguirre, D'Esposito, & Farah, 1999; Thompson-Schill et al., 1997; Wagner et al., 2001).

Looking first at the PTr, we suggest that speech-associated gestures are a source of semantic information that can be used by listeners to reduce the natural level of ambiguity
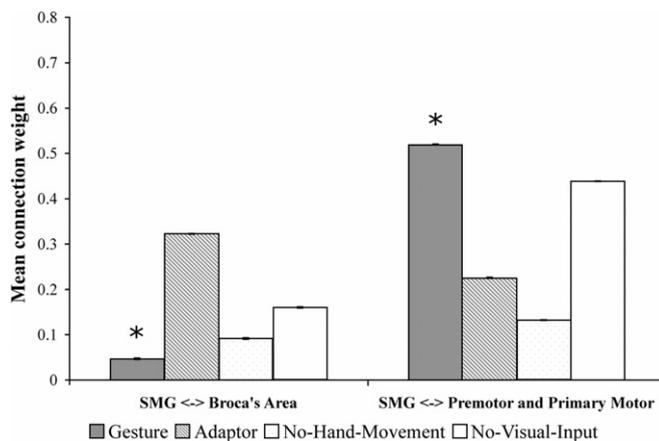

Fig. 5. Mean of all supramarginal gyrus (SMG) connection weights from Bayesian averaging from exhaustive search of all structural equation models shown in Fig. 4 with the pars opercularis and pars triangularis (i.e., Broca's area) and dorsal and ventral premotor and primary motor cortex. Asterisks indicate a significant difference between the *Gesture* condition and the *No-Hand-Movement*, *Self-Adaptor*, and *No-Visual-Input* conditions ($p < .00001$ in all cases). Error bars indicate standard error.
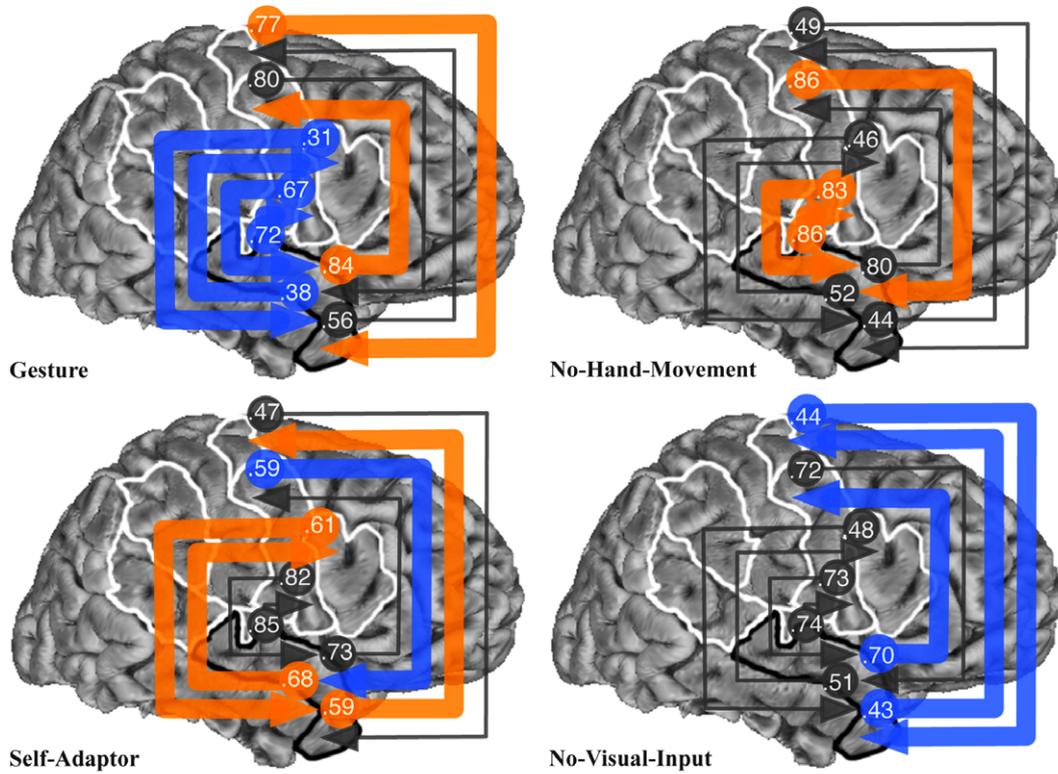
Fig. 6. Results of Bayesian averaging of connection weights from exhaustive search of all structural equation models with connections between STa and POp, PTr, PPMv, or PPMd regions of interest (ROIs), for the *Gesture*, *No-Hand-Movement*, *Self-Adaptor*, and *No-Visual-Input* conditions. See Fig. 1 caption for the definition of ROI abbreviations and Fig. 1 for location of ROIs. See Fig. 2 caption for the meaning of arrowed lines and their colors.
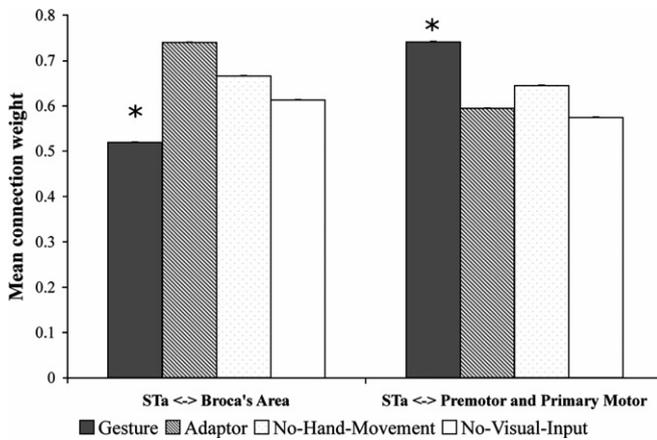

Fig. 7. Mean of all anterior superior temporal (STa) connection weights from Bayesian averaging from exhaustive search of all structural equation models shown in Fig. 6 with the pars opercularis and pars triangularis (i.e., Broca's area) and dorsal and ventral premotor and primary motor cortex. Asterisks indicate a significant difference between the *Gesture* condition and the *No-Hand-Movement*, *Self-Adaptor*, and *No-Visual-Input* conditions ($p < .00001$ in all cases). Error bars indicate standard error.

associated with spoken discourse. Activity levels in the PTr increase as the level of lexical ambiguity increases in sentences, presumably because there are increased retrieval or selection demands associated with ambiguous content (Rodd, Davis, & Johnsrude, 2005). The decreased involvement of the PTr when speech-associated gestures are present (and the increased involvement when they are not) is consistent with the interpretation that speech-associated gestures reduce ambiguity and, concomitantly, selection and retrieval needs.

Turning to the POp, we suggested in Section 2 that observing speech-associated gestures might reduce the influence of the POp on other areas because attention is directed at visual hand and arm movements. Because we are generally poor at dividing visual attention to objects, increased attention to the hands (and the message, or meaning, level of speech) should result in decreased attention to the face (and the phonological processing of speech). As a result, there will be less need to match visual information about mouth movements with motor plans for producing those movements. The relatively smaller influence of the POp on other areas during the *Gesture* condition (compared to the other conditions) supports this interpretation. Indeed, the *No-Hand-Movement* condition produced strong interactions between the POp and the STp, suggesting the use of mouth movements to aid in phonological aspects of speech perception. The *Self-Adaptor* condition produced strong influence of the POp (and the PTr) on other regions. It may be that these self-grooming movements are distracting, which then increases demands on phonological evaluation and retrieval/selection of semantic information.
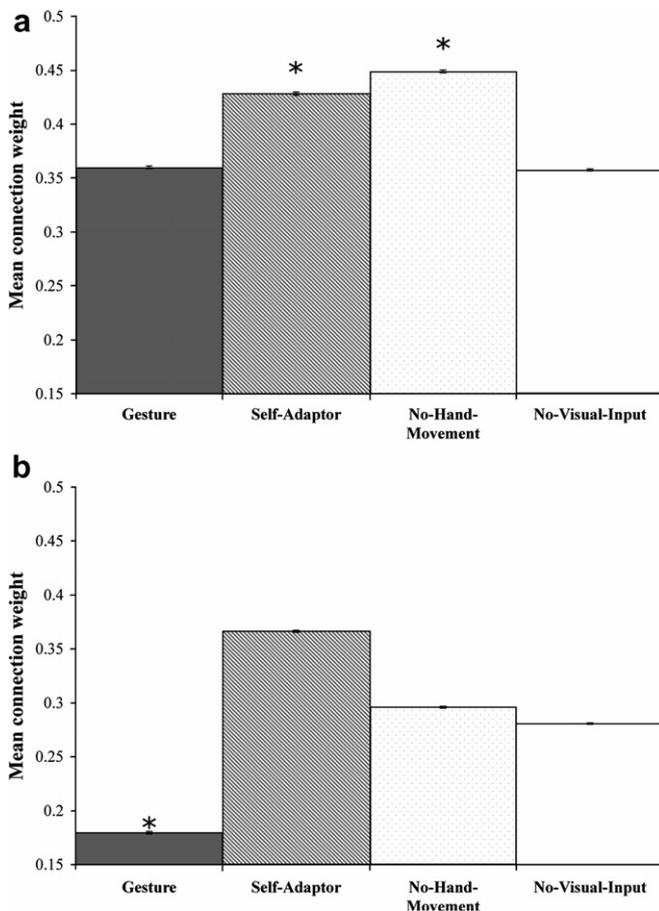
Fig. 8. Mean of all (a) pars opercularis (POp) and (b) pars triangularis (PTr) connection weights from Bayesian averaging from exhaustive search of all structural equation models shown in Figs. 2, 4, and 6. Asterisks indicate a significant difference between the *Gesture* condition and the *No-Hand-Movement*, *Self-Adaptor*, and/or *No-Visual-Input* conditions ($p < .00001$ in all cases). Error bars indicate standard error.

We also found that the reduction of PTr influence on other cortical areas in the *Gesture* condition (relative to the other conditions) was more profound than the reduction of POp influence. This difference may stem from the fact that listeners can shift attention from the message level to the phonological level at different points in discourse. For example, when there are no speech-associated gestures, attention may be directed at mouth movements. This information can then be used to reduce the ambiguity associated with speech sounds (see Section 2).

These interpretations are supported by a recent fMRI study showing increased activity in the PTr when listeners process gestures artificially constructed to mismatch the previous context established in speech (Willems, Özyürek, & Hagoort, 2006; note that this type of mismatch is different from naturally occurring gesture–speech mismatches in which gesture conveys information that is different from, but not necessarily contradictory to, the information conveyed in speech, Goldin-Meadow, 2003). By our interpretation (which is different from the authors'), gesture and speech are not co-expressive in the Willems et al. study,

and thus create an increased demand for semantic selection or retrieval. That is, like the *Self-Adaptor* condition, unnatural mismatching hand movements are distracting, which then increases demands with respect to retrieval/selection of semantic information.

Although these interpretations are speculative, we believe that they are more coherent than the view that Broca's area processes speech-associated gestures as part of the action recognition/production system. Given the proposed role of mirror neurons and the mirror system in processing goal-oriented actions, and the already demonstrated role of Broca's area in the observation and execution of face, finger, hand, and arm movements (e.g., Iacoboni et al., 1999), one would expect the POp and PTr to have the strongest influence on the rest of the brain when processing speech-associated gestures (which they did not). Similarly, given the hypothesis that the connection between action recognition and language functions attributed to Broca's area is the underlying goal of sequencing movements, one would expect the POp and PTr to play a relatively large role during the observation of speech-associated gestures (which they did not).

5.2. Speech-associated gestures and the human mirror system

The relative lack of involvement of Broca's area with other cortical areas in the presence of speech-associated gestures suggests that we should rethink the theoretical perspective in which the function of Broca's area is limited to action recognition or production. Indeed, it is possible that the mirror system in humans is a dynamically changing system (Arbib, 2006). In fact, a recent study shows that, when simple actions are viewed, activity within the mirror system is modulated by the motivation and goals of the perceiver (Cheng, Meltzoff, & Decety, 2006), an observation that goes well beyond the theoretical claims made to date about the mirror system. Thus, we suggest that what constitutes the human mirror system for action recognition may depend on the goal of the listener and the dynamic organization of the brain to accommodate that goal.

With respect to the present experiment, the goal of the listeners in all conditions was to comprehend spoken language. In different conditions, there were different visually perceptible actions that could be used (in addition to the auditory speech signal) to support the goal of comprehension. In the *Gesture* condition, speech-associated gestures were actions that could be recognized and used to reduce lexical or sentential ambiguity, with the result that comprehension was improved. In the *No-Hand-Movement* and *Self-Adaptor* conditions, face actions could be recognized and used to reduce ambiguity associated with speech, again with the result that comprehension was improved.

Rather than a mirror system that is static, our results suggest that the human mirror system (or, perhaps, mirror *systems*) dynamically changes to accommodate the goal of comprehension in response to the observed actions that are

available. When speech-associated gestures are present, there are strong interactions among the SMG, PPMv, PPMd, and STa (see the left-hand column of Table 2). When the available actions are face movements, there are strong interactions primarily among the POp, PPMv, STp, and STa (see the right hand column of Table 2).

We suggest that these differences in the strength of interactions among brain areas reflect the brain dynamically changing when different actions are relevant to the goal of language comprehension. Thus, when *speech-associated gestures* are observed, strong interactions between the SMG, PPMv, PPMd, and STa reflect the activity of a human mirror system because the SMG is involved in preparation for hand movements and both the PPMv and PPMd have representations used for producing hand movements. Interaction among these areas may constitute a mirror or observation–execution matching system for hand and arm movements. The strong interaction between the PPMv and PPMd and the STa, an area involved in semantic aspects of spoken language comprehension, may reflect the fact that recognition of these matched movements is relevant to language comprehension.

Similarly, when *face actions* are observed, strong interactions between the POp, PPMv, STp, and STa reflect the activity of a human mirror system because the POp is involved in matching acoustic and/or visual information about mouth movements with motor plans for producing those movements, and the PPMv is strongly involved in producing face and tongue movements (Fox et al., 2001; Schubotz & von Cramon, 2003). Interaction among these areas may constitute a mirror or observation–execution matching system for face movements. The strong interaction between the POp and PPMv and the STp, an area involved in, among other things, speech perception, could indicate the recognition of these observed and matched face actions as relevant to speech perception. Indeed, we have previously shown that the PPMv and STp are involved in using mouth movements to aid in the process of speech perception (Skipper et al., 2005, 2007). Strong interactions with the STa may occur because mouth movements may aid sentence comprehension when they help disambiguate the phonological properties of speech.

More generally, our results suggest that the human mirror system dynamically changes according to the observed action, and the relevance of that action, to understanding a given behavior. For example, if the behavioral goal involves understanding a speech sound when mouth movements can be observed, then areas of cortex involved in the execution of face movements and speech perception are likely to constitute the mirror system. In contrast, if the behavioral goal involves understanding a sentence when speech-associated gestures can be observed, then areas of cortex involved in the execution of hand movements and semantic aspects of language comprehension are likely to constitute the mirror system.

6. Conclusions

We interpret our results as suggesting that the interactions among the SMG, PPMv, PPMd, and STa reflect the activity of a human mirror system in extracting semantic information from speech-associated gestures, and integrating that information into the process of spoken language comprehension. Once this semantic information has been extracted and integrated, there is less need for phonological processing (associated with cortical interactions involving the POp) and less need for semantic retrieval and selection (associated with interactions involving the PTr). It has been argued that speech-associated gestures function primarily to aid lexical retrieval in speakers, rather than to communicate semantic content to listeners (Krauss, Dushay, Chen, & Rauscher, 1995; Krauss, Morrel-Samuels, & Colasante, 1991; Morsella & Krauss, 2004). However, our findings reinforce behavioral observations showing that speech-associated gestures do indeed function for listeners (as well as playing an important cognitive role for speakers, Goldin-Meadow, 2003)—speech-associated gestures communicate information that listeners are likely to use to reduce the need for lexical selection/retrieval. Because communication is a joint action of sorts (see Pickering & Garrod, 2004), information contained in speech-associated gestures is not likely to affect the speaker *or* the listener, but rather is likely to be mutually informative.

References

Alibali, M. W., Flevares, L. M., & Goldin-Meadow, S. (1997). Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology, 89*(1), 183–193.

Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H. B., & Zilles, K. (1999). Broca's region revisited: Cytoarchitecture and intersubject variability. *Journal of Comparative Neurology, 412*(2), 319–341.

Arbib, M. A. (2006). The mirror system hypothesis on the linkage of action and languages. In M. A. Arbib (Ed.), *Action to language via the mirror neuron system*. Cambridge, MA: Cambridge University Press.

Bates, E., Friederici, A., & Wulfeck, B. (1987). Comprehension in aphasia: A cross-linguistic study. *Brain and Language, 32*(1), 19–67.

Berger, K. W., & Popelka, G. R. (1971). Extra-facial gestures in relation to speech-reading. *Journal of Communication Disorders, 3*, 302–308.

Blank, S. C., Scott, S. K., Murphy, K., Warburton, E., & Wise, R. J. (2002). Speech production: Wernicke, Broca and beyond. *Brain, 125*(Pt. 8), 1829–1838.

Broca, M. P. (1861). Remarques sur le siége de la faculté du langage articulé, suivies d'une observation d'aphemie (Perte de la Parole). *Bulletins et Memoires de la Societe Anatomique de Paris, 36*, 330–357.

Brodmann, K. (1909). *Vergleichende localisationslehre der grosshirnrinde in ihren principien dargestellt auf grund des zellenbaues*. Leipzig: Barth.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience, 13*(2), 400–404.

Buccino, G., Binkofski, F., & Riggio, L. (2004). The mirror neuron system and action recognition. *Brain and Language, 89*(2), 370–376.

Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., Freund, H. J., et al. (2004). Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. *Neuron, 42*(2), 323–334.

Buchel, C., & Friston, K. J. (1997). Modulation of connectivity in visual pathways by attention: Cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex, 7*(8), 768–778.

Buchsbaum, B., Hickok, G., & Humphries, C. (2001). Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science, 25*, 663–678.

Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience, 12*(4), 679–690.

Cassell, J., McNeill, D., & McCullough, K. E. (1999). Speech-gesture mismatches: Evidence for one underlying representation of linguistic and non-linguistic Information. *Pragmatics and Cognitiion, 7*(1), 1–33.

Cheng, Y., Meltzoff, A. N., & Decety, J. (2006). Motivation modulates the activity of the human mirror-neuron system. *Cerebral Cortex.* doi:10.1093/cercor/bhl107.

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*(3), 162–173.

Crinion, J., & Price, C. J. (2005). Right anterior superior temporal activation predicts auditory sentence comprehension following aphasic stroke. *Brain, 128*(Pt. 12), 2858–2871.

Deacon, T. W. (1992). Cortical connections of the inferior arcuate sulcus cortex in the macaque brain. *Brain Research, 573*(1), 8–26.

Driskell, J. E., & Radtke, P. H. (2003). The effect of gesture on speech production and comprehension. *Human Factors, 45*(3), 445–454.

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature, 384*, 159–161.

Dronkers, N. (1998). Symposium: The role of Broca's area in language. *Brain and Language, 65*, 71–72.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1*(49–98).

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology, 73*(6), 2608–2611.

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage, 9*(2), 195–207.

Fischl, B., Sereno, M. I., Tootell, R. B., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping, 8*(4), 272–284.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D. H., et al. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex, 14*(1), 11–22.

Floel, A., Ellger, T., Breitenstein, C., & Knecht, S. (2003). Language perception activates the hand motor cortex: Implications for motor theories of speech perception. *European Journal of Neuroscience, 18*(3), 704–708.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine, 33*(5), 636–647.

Fox, P. T., Huang, A., Parsons, L. M., Xiong, J. H., Zamarippa, F., Rainey, L., et al. (2001). Location-probability profiles for the mouth region of human primary motor-sensory cortex: Model and validation. *Neuroimage, 13*(1), 196–209.

Fridman, E. A., Immisch, I., Hanakawa, T., Bohlhalter, S., Waldvogel, D., Kansaku, K., et al. (2006). The role of the dorsal stream for gesture production. *Neuroimage, 29*(2), 417–428.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*(Pt. 2), 593–609.

Gelfand, J. R., & Bookheimer, S. Y. (2003). Dissociating neural mechanisms of temporal sequencing and processing phonemes. *Neuron, 38*(5), 831–842.

Geschwind, N. (1965). The organization of language and the brain. *Science, 170*, 940–944.

Godschalk, M., Mitz, A. R., van Duin, B., & van der Burg, H. (1995). Somatotopy of monkey premotor cortex examined with microstimulation. *Neuroscience Research, 23*(3), 269–279.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

Goldin-Meadow, S., & Momeni Sandhofer, C. (1999). Gestures convey substantive information about a child's thoughts to ordinary listeners. *Developmental Science, 2*(1), 67–74.

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction, 9*(3), 201–219.

Gough, P. M., Nobre, A. C., & Devlin, J. T. (2005). Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. *Journal of Neuroscience, 25*(35), 8010–8016.

Grant, K.W., & Greenberg, S. (2001). *Speech intelligibility derived from asynchronous processing of auditory-visual information*. Paper presented at the Proceedings of the Workshop on Audio-Visual Speech Processing, Scheelsminde, Denmark.

Grezes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurones in the human brain: An fMRI study. *Neuroimage, 18*(4), 928–937.

Hamzei, F., Rijntjes, M., Dettmers, C., Glauche, V., Weiller, C., & Buchel, C. (2003). The human action recognition system and its relationship to Broca's area: An fMRI study. *Neuroimage, 19*(3), 637–644.

Hanson, S. J., Matsuka, T., Hanson, C., Rebbechi, D., Halchenko, Y., Zaimi, A., et al. (2004). *Structural equation modeling of neuroimaging data: Exhaustive search and markov chain Monte Carlo*. Paper presented at the Organization for Human Brain Mapping, Budapest, Hungry.

Harrington, D. L., Rao, S. M., Haaland, K. Y., Bobholz, J. A., Mayer, A. R., Binderx, J. R., et al. (2000). Specialized neural systems underlying representations of sequential movements. *Journal of Cognitive Neuroscience, 12*(1), 56–77.

Hasson, U., Skipper, J. I., Wilde, M. J., Nusbaum, H. C., & Small, S. L. (submitted for publication). Improving the analysis, storage and sharing of imaging data using relational databases and distributed computing. *Neuroimage*.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition, 92*(1–2), 67–99.

Hluštík, P., Solodkin, A., Skipper, J. I., & Small, S. L. (in preparation). Somatotopic mapping of dorsal and ventral premotor cortex using sequential pointing with the eyes, finger, foot, and tongue.

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial. *Statistical Science, 14*(4), 382–417.

Holler, J., & Beattie, G. (2003). Pragmatic aspects of representational gestures: Do speakers use them to clarify verbal ambiguity for the listener? *Gesture, 3*(2), 127–154.

Horwitz, B., Tagamets, M. A., & McIntosh, A. R. (1999). Neural modeling, functional brain imaging, and cognition. *Trends in Cognitive Sciences, 3*(3), 91–98.

Humphries, C., Love, T., Swinney, D., & Hickok, G. (2005). Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Human Brain Mapping*.

Humphries, C., Willard, K., Buchsbaum, B., & Hickok, G. (2001). Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. *Neuroreport, 12*(8), 1749–1752.

Iacoboni, M. (2005). Understanding others: Imitation, language, empathy. In: S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From cognitive neuroscience to social science, Vol. 1: Mechanisms of Imitation and Imitation in Animals* (pp. 77–99). Cambridge, MA: MIT Press.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286*(5449), 2526–2528.

Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics, 5*(3), 299–314.

Josse, G., Dick, A. S., Skipper, J. I., Nusbaum, H. C., Goldin-Meadow, S., & Small, S. L. (in preparation). Speech-associated gestures affect neural activity in cortical areas associated with semantic processing.

Josse, G., Skipper, J. I., Chen, E. E., Goldin-Meadow, S., & Small, S. L. (2006). *Meaningful mouth and hand movements play a role in language comprehension: Evidence from fMRI*. Paper presented at the Organization for Human Brain Mapping.

Josse, G., Suriyakham, L. W., Skipper, J. I., Levine, S. C., Goldin-Meadow, S., Huttenlocher, P., et al. (2005). *Language-associated gesture modulates areas of the brain involved in action observation and language comprehension*. Paper presented at the Organization for Human Brain Mapping, Toronto, Canada.

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science, 274*(5284), 114–116.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of The American Statistical Association, 90*, 773–795.

Kendon, A. (1987). On gesture: Its complementary relationship with speech. In A. Siegman & S. Feldstein (Eds.), *Nonverbal communication* (pp. 65–97). Hillsdale, New Jersey: Lawrence Erlbaum.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge, England: Cambridge University Press.

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences*. Pacific Grove, CA: Brooks-Cole Publishing Company.

Knopman, D. S., Selnes, O. A., Niccum, N., Rubens, A. B., Yock, D., & Larson, D. (1983). A longitudinal study of speech fluency in aphasia: CT correlates of recovery and persistent nonfluency. *Neurology, 33*(9), 1170–1178.

Kohler, E., Keysers, C., Umilta, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science, 297*(5582), 846–848.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture*. Cambridge, UK: Cambridge University Press.

Krauss, R. M., Dushay, R., Chen, Y., & Rauscher, F. (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology, 31*, 533–552.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology, 61*, 743–754.

Madigan, D., & Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *Journal of The American Statistical Association, 89*, 1535–1546.

McIntosh, A. R., & Gonzalez-Lima, F. (1991). Structural modeling of functional neural pathways mapped with 2-deoxyglucose: Effects of acoustic startle habituation on the auditory system. *Brain Research, 547*(2), 295–302.

McIntosh, A. R., & Gonzalez-Lima, F. (1992). Structural modeling of functional visual pathways mapped with 2-deoxyglucose: Effects of patterned light and footshock. *Brain Research, 578*(1–2), 75–86.

McIntosh, A. R., & Gonzalez-Lima, F. (1993). Network analysis of functional auditory pathways mapped with fluorodeoxyglucose: Associative effects of a tone conditioned as a Pavlovian excitor or inhibitor. *Brain Research, 627*(1), 129–140.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: The Universitty of Chicago Press.

McNeill, D., Cassell, J., & McCullough, K. E. (1994). Communicative effects of speech-mismatched gestures. *Research on Language and Social Interaction, 27*, 223–237.

Mohr, J. P., Pessin, M. S., Finkelstein, S., Funkenstein, H. H., Duncan, G. W., & Davis, K. R. (1978). Broca aphasia: Pathologic and clinical. *Neurology, 28*(4), 311–324.

Molnar-Szakacs, I., Iacoboni, M., Koski, L., Maeda, F., Dubeau, M.C., Aziz-Zadeh, L., et al. (2002). *Action observation in the pars opercularis: Evidence from 58 subjects studied with fMRI*. Paper presented at the Cognitive Neuroscience Society, San Francisco.

Morsella, E., & Krauss, R. M. (2004). The role of gestures in spatial working memory and speech. *American Journal of Psychology, 117*(3), 411–424.

Moss, H. E., Abdallah, S., Fletcher, P., Bright, P., Pilgrim, L., Acres, K., et al. (2005). Selecting among competing alternatives: Selection and retrieval in the left inferior frontal gyrus. *Cerebral Cortex, 15*(11), 1723–1735.

Mottonen, R., Jarvelainen, J., Sams, M., & Hari, R. (2005). Viewing speech modulates activity in the left SI mouth cortex. *Neuroimage, 24*(3), 731–737.

Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G., & Orban, G. A. (2005). Observing others: Multiple action representation in the frontal lobe. *Science, 310*(5746), 332–336.

Nishitani, N., & Hari, R. (2002). Viewing lip forms: Cortical dynamics. *Neuron, 36*(6), 1211–1220.

Nishitani, N., Schurmann, M., Amunts, K., & Hari, R. (2005). Broca's region: From action to language. *Physiology (Bethesda), 20*, 60–69.

Noll, D. C., Cohen, J. D., Meyer, C. H., & Schneider, W. (1995). Spiral k-space MRI imaging of cortical activation. *Journal of Magnetic Resonance Imaging, 5*, 49–56.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113.

Penny, W. D., Stephan, K. E., Mechelli, A., & Friston, K. J. (2004). Modelling functional integration: A comparison of structural equation and dynamic causal models. *Neuroimage, 23*(Suppl. 1), S264–S274.

Petrides, M., & Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience, 16*(2), 291–310.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*(2), 169–190, discussion 190–226.

Poldrack, R. A., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenich, M., et al. (2001). Relations between the neural bases of dynamic auditory processing and phonological processing: Evidence from fMRI. *Journal of Cognitive Neuroscience, 13*(5), 687–697.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage, 10*(1), 15–35.

Records, N. L. (1994). A measure of the contribution of a gesture to the perception of speech in listeners with aphasia. *Journal of Speech and Hearing Research, 37*(5), 1086–1099.

Reisberg, D., McLean, J., & Goldfield, A. (1987). Easy to hear but hard to understand: A lipreading advantage with intact auditory stimuli. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lipreading* (pp. 97–114). Hillsdale, NJ: Erlbaum.

Rimé, B. (1982). The elimination of visible behavior from social interactions: Effects on verbal, nonverbal and interpersonal variables. *European Journal of Social Psychology, 12*, 113–129.

Risberg, A., & Lubker, J. (1978). Prosody and speechreading. *Speech Transmission Laboratory Quarterly Progress Report Status Report, 4*, 1–16.

Riseborough, M. G. (1981). Physiographic gestures as decoding facilitators: Three experiments exploring a neglected facet of communication. *Journal of Nonverbal Behavior, 5*, 172–183.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences, 21*(5), 188–194.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169–192.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Research. Cognitive Brain Research, 3*(2), 131–141.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Current Opinion in Neurobiology, 12*(2), 149–154.

Rizzolatti, G., Luppino, G., & Matelli, M. (1998). The organization of the cortical motor system: New concepts. *Electroencephalography and Clinical Neurophysiology, 106*(4), 283–296.

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex, 15*(8), 1261–1269.

Rogers, W. T. (1978). The contribution of kinesic illustrators toward the comprehension of verbal behavior within utterances. *Human Communication Research, 5*, 54–62.

Rowe, J., Friston, K., Frackowiak, R., & Passingham, R. (2002). Attention to action: Specific modulation of corticocortical interactions in humans. *Neuroimage, 17*(2), 988–998.

Schubotz, R. I., & von Cramon, D. Y. (2003). Functional-anatomical concepts of human premotor cortex: Evidence from fMRI and PET studies. *Neuroimage, 20*(Suppl. 1), S120–S131.

Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science, 16*(2), 85–89.

Skipper, J. I., Goldin-Meadow, S., Nusbaum, H.C., & Small, S.L. (in preparation). Gestures in the brain's language.

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: Motor cortical activation during speech perception. *Neuroimage, 25*(1), 76–89.

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2006). Lending a helping hand to hearing: Another motor theory of speech perception. In M. A. Arbib (Ed.), *Action to language via the mirror neuron system*. Cambridge, MA: Cambridge University Press.

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex*. doi:10.1093/cercor/bhl147.

Skipper, J. I., Nusbaum, H. C., van Wassenhove, V., Barber, C., Chen, E. E., & Small, S.L. (2005). *The role of ventral premotor and primary motor cortex in audiovisual speech perception*. Paper presented at the Organization for Human Brain Mapping, Toronto, Canada.

Solodkin, A., Hlustik, P., Chen, E. E., & Small, S. L. (2004). Fine modulation in network activation during motor execution and motor imagery. *Cerebral Cortex, 14*(11), 1246–1255.

Sumby, W. H., & Pollack, I. (1954). Visual contribution of speech intelligibility in noise. *The Journal of the Acoustical Society of America, 26*(2), 212–215.

Sundara, M., Namasivayam, A. K., & Chen, R. (2001). Observation-execution matching system for speech: A magnetic stimulation study. *Neuroreport, 12*(7), 1341–1344.

Thompson, L. A., & Massaro, D. W. (1986). Evaluation and integration of speech and pointing gestures during referential understanding. *Journal of Experimental Child Psychology, 42*(1), 144–168.

Thompson-Schill, S. L., Aguirre, G. K., D'Esposito, M., & Farah, M. J. (1999). A neural basis for category and modality specificity of semantic knowledge. *Neuropsychologia, 37*(6), 671–676.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America, 94*(26), 14792–14797.

Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience, 14*(4), 550–560.

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., et al. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage, 30*(4), 1414–1432.

Wagner, A. D., Pare-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron, 31*(2), 329–338.

Watkins, K., & Paus, T. (2004). Modulation of motor excitability during speech perception: The role of Broca's area. *Journal of Cognitive Neuroscience, 16*(6), 978–987.

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia, 41*(8), 989–994.

Willems, R. M., Özyürek, A., & Hagoort, P. (2006). When language meets action: The neural integration of gesture and speech. *Cerebral Cortex*. doi:10.1093/cercor/bhl141.

Wise, R. J., Greene, J., Buchel, C., & Scott, S. K. (1999). Brain regions involved in articulation. *Lancet, 353*(9158), 1057–1061.