

Symbolic gestures and spoken language are processed by a common neural system

Jiang Xu^a, Patrick J. Gannon^b, Karen Emmorey^c, Jason F. Smith^d, and Allen R. Braun^{a,1}

^aLanguage Section, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Building 10, Room 8S235A, Bethesda, MD 20892; ^bDepartment of Science Education, Hofstra University School of Medicine, North Shore Long Island Jewish Health, Hempstead, NY 11549; ^cLaboratory for Language and Cognitive Neuroscience, San Diego State University, 6495 Alvarado Road, Suite 200, San Diego, CA 92120; and ^dBrain Imaging and Modeling Section, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Building 10, Room 8S235B, Bethesda, MD 20892

Edited by Leslie G. Ungerleider, National Institutes of Health, Bethesda, MD, and approved October 6, 2009 (received for review August 14, 2009)

Symbolic gestures, such as pantomimes that signify actions (e.g., threading a needle) or emblems that facilitate social transactions (e.g., finger to lips indicating “be quiet”), play an important role in human communication. They are autonomous, can fully take the place of words, and function as complete utterances in their own right. The relationship between these gestures and spoken language remains unclear. We used functional MRI to investigate whether these two forms of communication are processed by the same system in the human brain. Responses to symbolic gestures, to their spoken glosses (expressing the gestures’ meaning in English), and to visually and acoustically matched control stimuli were compared in a randomized block design. General Linear Models (GLM) contrasts identified shared and unique activations and functional connectivity analyses delineated regional interactions associated with each condition. Results support a model in which bilateral modality-specific areas in superior and inferior temporal cortices extract salient features from vocal-auditory and gestural-visual stimuli respectively. However, both classes of stimuli activate a common, left-lateralized network of inferior frontal and posterior temporal regions in which symbolic gestures and spoken words may be mapped onto common, corresponding conceptual representations. We suggest that these anterior and posterior perisylvian areas, identified since the mid-19th century as the core of the brain’s language system, are not in fact committed to language processing, but may function as a modality-independent semiotic system that plays a broader role in human communication, linking meaning with symbols whether these are words, gestures, images, sounds, or objects.

evolution | fMRI | perisylvian | communication | modality-independent

Symbolic gestures, found in all human cultures, encode meaning in conventionalized movements of the hands. Commonplace gestures such as those depicted in Fig. 1 are routinely used to facilitate social transactions, convey requests, signify objects or actions, or comment on the actions of others. As such, they play a central role in human communication.

The relationship between symbolic gestures and spoken language has been the subject of longstanding debate. A key question has been whether these represent parallel cognitive domains that may specify the same semantic content but intersect minimally (1, 2), or whether, at the deepest level, they reflect a common cognitive architecture (3, 4). In support of the latter idea, it has been argued that words can be viewed as arbitrary and conventionalized vocal tract gestures that encode meaning in the same way that gestures of the hands do (5). In this view, symbolic gesture and spoken language might be best understood in a broader context, as two parts of a single entity, a more fundamental semiotic system that underlies human discourse (3).

The issue can be reframed in anatomical terms, providing a testable set of questions that can be addressed using functional neuroimaging methods: are symbolic gestures and language processed by the same system in the human brain? Specifically, does the so called “language network” in the inferior frontal and posterior temporal cortices surrounding the sylvian fissure support the pair-



Fig. 1. Examples of pantomimes (top two rows. English glosses—A. juggle balls; B. unscrew jar) and emblems (bottom two rows. English glosses—C. I’ve got it!; D. settle down).

ing of form and meaning for both forms of communication? We used functional MRI (fMRI) to address these questions.

Symbolic gestures constitute a subset within in the wider human gestural repertoire. They are first of all voluntary, and thus do not include the large class of involuntary gestures—hardwired, often

Author contributions: J.X., P.J.G., K.E., and A.R.B. designed research; J.X. performed research; J.F.S. contributed new reagents/analytic tools; J.X. and A.R.B. analyzed data; and J.X., P.J.G., K.E., and A.R.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. E-mail: brauna@nidcd.nih.gov.

This article contains supporting information online at www.pnas.org/cgi/content/full/0909197106/DCSupplemental.

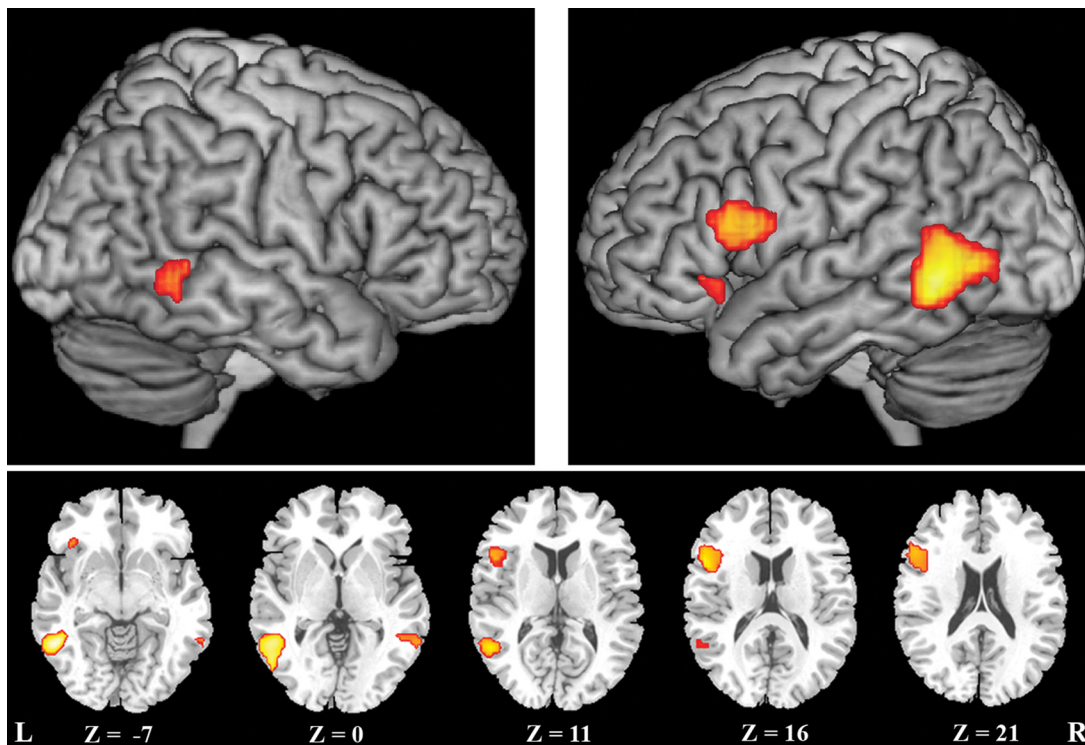


Fig. 2. Common areas of activation for processing symbolic gestures and spoken language minus their respective baselines, identified using a random effects conjunction analysis. The resultant *t* map is rendered on a single subject T1 image: 3D surface rendering above, axial slices with associated *z* axis coordinates, below.

visceral responses that indicate pleasure, anger, fear, disgust (4). Of the voluntary gestures described by McNeill (3), symbolic gestures also form a discrete subset. Symbolic gestures are conventionalized, conform to a set of cultural standards and therefore do not include gesticulations that accompany speech (3, 4). Speech-related gestures are idiosyncratic, shaped by each speaker's own communicative needs, while symbolic gestures are learned, socially transmitted, and form a common vocabulary (see [supporting information \(SI\) Note 1](#)). Crucially, while they amplify the meaning conveyed in spoken language, speech-related gestures do not encode and communicate meaning in themselves in the same way as symbolic gestures, which can fully take the place of words and function as complete utterances in their own right (i.e., as independent, self-contained units of discourse).

Which symbolic gestures can be used to address our central question? For many reasons, sign languages are not ideal for this purpose. Over the past decade, neuroimaging studies have clearly and reproducibly demonstrated that American Sign Language, British Sign Language, Langue des Signes Québécoise, indeed all sign languages studied thus far, elicit patterns of activity in core perisylvian areas that are, for the most part, indistinguishable from those accompanying the production and comprehension of spoken language (6). But this is unsurprising. These are all natural languages, which by definition communicate propositions by means of a formal linguistic structure, conforming to a set of phonological, lexical, and syntactic rules comparable to those that govern spoken or written language. The unanimous interpretation, reinforced by the sign aphasia literature (6), has been that the perisylvian cortices process languages that possess this canonical, rule-based structure, independent of the modality in which they are expressed.

But demonstrating the brain's congruent responses to signed and spoken language cannot answer our central question, that is, whether or not the perisylvian areas play a broader role in mediating symbolic—nonlinguistic as well as linguistic—communication. To do so, it will be necessary to systematically exclude spoken, written,

or signed language, minimize any potential overlap and consider autonomous nonlinguistic gestures.

Human gestures, idiosyncratic as well as conventionalized, have been ordered in a useful way by McNeill (3) according to what he has termed “Kendon’s continuum.” This places gestural categories in a linear sequence—gesticulations to pantomimes to emblems to sign languages—according to a decreasing presence of speech and increasing language-like features. As noted, although they appear at opposite ends of this continuum, cospeech gestures and sign languages are both fundamentally related to language, cannot be experimentally dissociated from it, and cannot be used to unambiguously address the questions we pose.

In contrast, the gestures at the core of the continuum, pantomimes and emblems, are ideal for this purpose. While their surface attributes differ—pantomimes directly depict or mimic actions or objects in themselves while emblems signify more abstract propositions (4)—the essential features that they share are the most important: They are semiotic gestures and they use movements of hands to symbolically encode and communicate meaning, but, crucially, they do so independent of language. They are autonomous and do not have to be related to—or recoded into—language to be understood.

Of the two, pantomimes have been studied more frequently using neuroimaging methods. A number of these studies have provided important information about the surface features of this gestural class, evaluated in the context of praxis (7), or in studies of action observation or imitation (8), often to assess the role of the mirror neuron system in pantomime perception. To date, however, no functional imaging study has directly compared comprehension of spoken language and pantomimes that convey identical semantic information. Similarly, emblems have been investigated in a variety of ways—by modifying their affective content, (9, 10) comparing them to other gestural classes or to baseline conditions (8, 11)—but not by contrasting them directly with spoken glosses that communicate the same information.

A number of studies have considered how emblems and spoken language reinforce or attenuate one another, or how these gestures modulate the neural responses to speech (e.g., ref. 12). Our goal is distinct in that rather than focusing on interactions between symbolic gesture and speech, we consider how the brain processes these independently, comparing responses to nonlinguistic symbolic gestures with those elicited by their spoken English glosses, to determine whether and to what degree these processes overlap.

The first possibility—that there is a unique, privileged portion of the brain solely dedicated to the processing of language—leads to a clear prediction: when functional responses to symbolic gestures and equivalent sets of words are compared, the functional overlap should be minimal. The alternative—that the inferior frontal and posterior temporal cortices do not constitute a language network *per se* but function as a general, modality-independent system that supports symbolic communication—predicts instead that the functional overlap should be substantial. We hypothesized that we would find the latter and that the shared features would be interpretable in light of what is known about how the brain processes semantic information.

Results

Blood oxygenation level-dependent (BOLD) responses to pantomimes and emblems, their spoken glosses, and visually and acoustically matched control tasks were compared in a randomized block design. Task recognition tests showed that subjects accurately identified $75 \pm 7\%$ of stimulus items, indicating that they had attended throughout the course of the experiment. Contrast analyses identified shared activations (conjunctions, seen for both gestures and the spoken English glosses), as well as activations that were unique to each condition. Seed region connectivity analyses were also performed to detect functional interactions associated with these conditions.

Conjunction and Contrast Analyses. Common activations for symbolic gesture and spoken language. Random effects conjunctions (conjunction null, SPM2) between weighted contrasts [(pantomime gestures + emblem gestures) – nonsense gesture controls] and [(pantomime glosses + emblem glosses) – pseudosound controls] revealed a similar pattern of responses evoked by gesture and spoken language in posterior temporal and inferior frontal regions (Fig. 2, Table 1).

Common activations in the inferior frontal gyrus (IFG) were found only in the left hemisphere. These included a large cluster encompassing the pars opercularis (BA44) and pars triangularis (BA45) and a discrete cluster in the pars orbitalis (BA47) (Fig. 2, Table 1). While activations in the posterior temporal lobe were bilateral, they were more robust, with a larger spatial extent, in the left hemisphere. There they were maximal in the posterior middle temporal gyrus (pMTG, BA21), extending dorsally through the posterior superior temporal sulcus (STS, BA21/22) toward the superior temporal gyrus (STG) and ventrally into the dorsal bank of the inferior temporal sulcus. (Fig. 2). In the right hemisphere activations were significant in the pMTG. (Fig. 2, Table 1).

Unique Activations for Symbolic Gesture and Spoken Language. Random effects subtractions of the same weighted contrasts showed that symbolic gestures, but not speech, elicited significant activations ventral to the areas identified by the conjunction analysis, which extended from the fusiform gyrus dorsally and laterally to include inferior temporal and occipital gyri in both hemispheres. These activations were maximal in the left inferior temporal (ITG, BA20) and in the right fusiform gyrus (BA37) (Fig. 3, Table 1).

These contrasts also revealed bilateral activations that were elicited by spoken language but not gesture, anterior to the areas identified by the conjunction analysis. These were maximal in the MTG (BA21), extending dorsally into anterior portions of the STS (BA21/22) (Fig. 3, Table 1) and STG (BA22) in both hemispheres.

Table 1. Brain responses to symbolic gestures and spoken language

Region	BA	Cluster	x	y	z	t-score	P
Conjunctions							
L post MTG	21	322	-51	-51	-6	4.49	0.0001†
L post STS	21/22		-54	-54	12	3.25	0.001†
L dorsal IFG	44/45	150	-54	12	15	3.34	0.001†
L ventral IFG	47	19	-39	27	-3	2.73	0.005
R post MTG	21	45	60	-51	-3	2.92	0.003
Speech > gesture							
L anterior MTG	21	507	-66	-24	-9	6.08	0.0001†
L anterior STS	21/22		-60	-36	1	5.24	0.0001†
R anterior MTG	21	417	60	-33	-6	6.25	0.0001†
R anterior STS	21/22		57	-36	2	5.35	0.0001†
Gesture > speech							
L fusiform/ITG	20/37	34	-45	-66	-12	3.46	0.001
R fusiform/ITG	20/37	75	45	-63	-15	3.42	0.001

Conjunctions and condition-specific activations. Regions of interest, Brodmann numbers, and MNI coordinates indicating local maxima of significant activations are tabulated with associated t-scores, probabilities, and cluster sizes. †Indicates $P < 0.05$, FDR corrected.

Connectivity Analyses. Regressions between BOLD signal variations in seed regions of interest (left IFG and pMTG) and other brain areas were derived independently for gesture and speech conditions and then statistically compared. These comparisons yielded maps that were used to identify interregional correlations common to gesture and speech and those that were unique to each condition. Common correlations were more abundant than those that were condition-dependent.

Common Connectivity Patterns for Symbolic Gesture and Spoken Language. Significant associations present for both gesture and speech are depicted in Fig. 4 (see also Table S1). These include strong functional connections between the IFG and pMTG in the left hemisphere and between each of these seed regions and their homologues on the right. Both seed regions were also coupled to extrasyllabic areas, including dorsolateral prefrontal cortex bordering on the lateral premotor cortices (BA9/6), parahippocampal gyri and contiguous mesial temporal structures (BA35 and 36) in both hemispheres, and the right cerebellar crus.

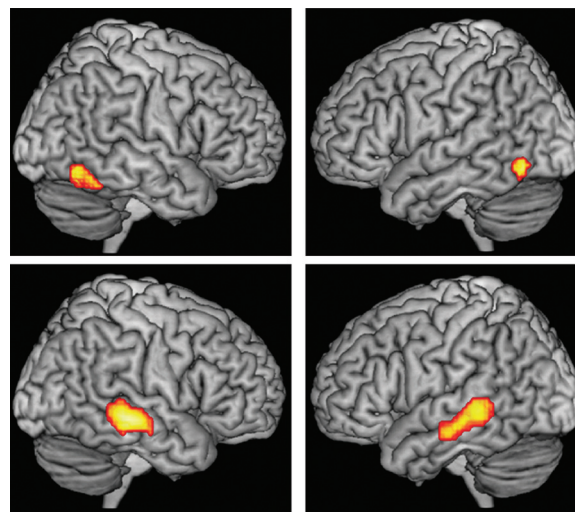


Fig. 3. Condition-specific activations for symbolic gesture (Top) and for speech (Bottom) were estimated by contrasting symbolic gestures and spoken language (minus their respective baselines) using random effects, paired two-sample t tests. Maps were rendered as indicated as in Fig. 2.

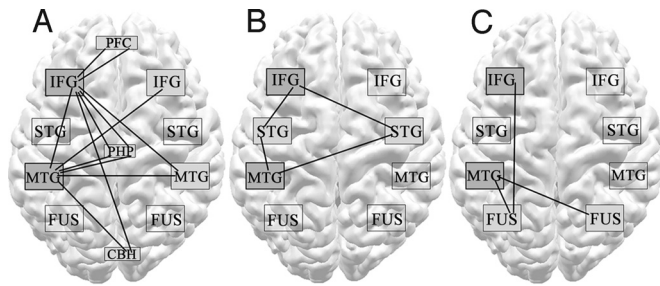


Fig. 4. Functional connections between left hemisphere seed regions (IFG; MTG) and other brain areas. (A) Correlations that exceeded threshold for both symbolic gesture and spoken language and did not differ significantly between these conditions. Unique correlations, illustrated in (B) for speech and (C) for gesture, exceeded threshold in only one of these conditions. Lines connote z-transformed correlation coefficients >3 that met criteria. (STG, superior temporal gyrus; FUS, inferior temporal and fusiform gyri; PFC, dorsolateral prefrontal cortex; PHP, parahippocampal gyrus; CBH, cerebellar hemisphere).

Unique Connectivity Patterns for Symbolic Gesture and Spoken Language.

In the gesture condition alone, activity in the left IFG was significantly correlated with that in the left ventral temporal cortices (a cluster encompassing fusiform and ITG); the left MTG was significantly correlated with activity in fusiform and ITG in both left and right hemispheres (Fig. 4, Table S1).

During the speech condition alone, both IFG and pMTG seed regions were functionally connected to the left STS and the STG bilaterally (Fig. 4, Table S1).

Some unique connections were found at the margins of clusters that were common to both gesture and speech (reported above). Thus, the spatial extent of IFG correlations with the right cerebellar crus was greater for gesture than for speech. Correlations between the IFG seed and more dorsal portions of the left IFG were larger for symbolic gesture; both IFG and pMTG seeds were, on the other hand, more strongly coupled to the left ventral IFG for speech.

Discussion

In order to understand to what degree systems that process symbolic gesture and language overlap, we compared the brain’s responses to emblems and pantomimes with responses elicited by spoken English glosses that conveyed the same information. We used an additional set of tasks that were devoid of meaning but were designed to control for surface level features—from lower level visual and auditory elements to more complex properties such as human form and biological motion—to focus on the level of semantics.

Given the limited spatial and temporal resolution of the method, BOLD fMRI results need to be interpreted cautiously (see SI Note 2). Nevertheless, our data suggest that comprehension of both forms of communication is supported by a common, largely overlapping network of brain regions. The conjunctions—regions activated during processing of meaningful information conveyed in either modality—were found in the network of inferior frontal and posterior temporal regions located along the sylvian fissure that has been regarded since the mid-19th century as the core of the language system in the human brain. Nevertheless, the results indicate that rather than representing a system principally devoted to the processing of language, these regions may constitute the nodes of a broader, domain-general network that supports symbolic communication.

Shared Activations for Perception of Symbolic Gesture and Spoken Language. Conjunctions included the frontal opercular regions historically referred to as “Broca’s area.” The IFG plays an unambiguous role in language processing and damage to the area

typically results in spoken and sign language aphasia (6, 13, 14). Neuroimaging studies in healthy subjects show that the IFG is reliably activated during the processing of spoken, written, and signed language at the levels we have focused upon here, i.e., comprehension of words or of simple syntactic structures (6, 13, 14). While some neuroimaging studies of symbolic gesture recognition have not reported inferior frontal activation (9, 10), many others have detected significant responses for observation of either pantomimes (15–17) or emblems (8, 11, 17).

Conjunctions also included the posterior MTG and STS, central elements in the collection of regions traditionally termed “Wernicke’s area.” Like the IFG, activation of these regions has been demonstrated consistently in neuroimaging studies of language comprehension, for listening as well as reading, at the level of words and simple argument structures examined here (13, 14). Similar patterns of activation are also characteristically observed for comprehension of signed language (6) (see SI Note 3), and lesions of these regions typically result in aphasia for both spoken and signed languages (6, 18). In addition to activation during language tasks, the same posterior temporal regions have been shown to respond when subjects observe both pantomimes (8, 15, 17) and emblems (9, 11, 17), although it should be noted that some studies of symbolic gesture recognition have not reported activation of either the pMTG or STS (10).

Perisylvian Cortices as a Domain-General Semantic Network. What these regions have in common, consistent with their responses to gestures as well as spoken language, is that each appears to play a role in processing semantic information in multiple sensory modalities and across cognitive domains.

A number of studies have provided evidence that the IFG is activated during tasks that require retrieval of lexical semantic information (e.g., ref. 19) and may play a more domain-general role in rule-based selection (20), retrieval of meaningful items from a field of competing alternatives (21), and subsequent unification or binding processes in which items are integrated into a broader semantic context (22).

Similarly, neuroimaging studies have shown activation of the posterior temporal cortices during tasks that require content-specific manipulation of semantic information (e.g., ref. 23). It has been proposed that the pMTG is the site where lexical representations may be stored (24, 25), but there is evidence that it is engaged in more general manipulation of nonlinguistic stimuli as well (e.g., ref. 23). Clinical studies also support the idea that the posterior temporal cortices participate in more domain-general semantic operations: patients with lesions in the pMTG (in precisely the area highlighted by the conjunction analysis) present with aphasia, as expected, but also have significant deficits in nonlinguistic semantic tasks, such as picture matching (26). Similarly, auditory agnosia for nonlinguistic stimuli frequently accompanies aphasia in patients with posterior temporal damage (27).

The posterior portion of the STS included in the conjunctions, the so-called STS-ms, is a focal point for cross-modal sensory associations (28) and may play a role in integration of linguistic and paralinguistic information encoded in auditory and visual modalities (29, 30). For this reason, it is not surprising that the region responds to meaningful gestural-visual as well as vocal-auditory stimuli.

Clearly, the inferior frontal and posterior temporal areas do not operate in isolation but function collectively as elements of a larger scale network. It has been proposed that the posterior middle temporal and inferior frontal gyri may constitute the central nodes in a functionally-defined cortical network for semantics (31). This model, drawing upon others (22), may provide a more detailed framework for interpretation of our results. During language processing, according to Lau et al. (31), this fronto-temporal network enables access to lexical representations from long-term memory and integration of these into an ongoing semantic context.

According to this model, the region encompassing the left pMTG (and the contiguous STS and ITG) is the site of long-term storage of lexical information (the so-called lexical interface). When candidate representations are activated within these regions, retrieval of suitable items is guided by the IFG, consistent with its role in rule-based selection and semantic integration.

An earlier, related model accounted for mapping of sound and meaning in the processing of spoken language (32). Lau et al. (31) have expanded this to accommodate written language, and their model is thus able to account for the processing of lexical information independent of input modality. In this sense, the inferior frontal and posterior temporal cortices constitute an amodal network for the storage and retrieval of lexical semantic information. But the role played by this system may be more inclusive. Lau et al. (31) suggest that the posterior temporal regions, rather than simply storing lexical representations, could store higher level conceptual features, and they point out that a wide range of nonlinguistic stimuli evoke responses markedly similar to the N400, the source of which they argue is the pMTG. [In light of these arguments, it is useful to note that a robust N400 is also elicited by gestural stimuli (e.g., ref. 33)].

Our results support a broader role for this system, indicating that the perisylvian network may map symbols and their associated meanings in a more universal sense, not limited to spoken or written language, to sign language, or to language per se. That is, for spoken language, the system maps sound and meaning in the vocal-auditory domain. Our results suggest that it performs the same transformations in the gestural-visual domain, enabling contact between gestures and the meanings they encode (see *SI Note 4*).

Contrast and Connectivity Analyses. The hypothesis that the inferior frontal and posterior temporal cortices play a broader, domain-general role in symbolic communication is supported by the connectivity analyses, which revealed overlapping patterns of connections between each of the frontal and temporal seed regions and other areas of the brain during processing of both spoken language and symbolic gesture. These analyses demonstrated integration within the perisylvian system during both conditions (that is the left pMTG and IFG were functionally coupled to one another in each instance, as well as to the homologues of these regions in the right hemisphere). In addition, during both conditions, the perisylvian seed regions were coupled to extrasylvian areas including dorso-lateral prefrontal cortex, cerebellum, parahippocampal gyrus and contiguous mesial temporal areas, which may play a role in executive control and sensorimotor or mnemonic processes associated with the processing of symbolic gesture as well as speech.

However, both task contrast and connectivity analyses also identified patterns apparent only during the processing of symbolic gesture or speech. Contrast analyses revealed bilateral activation of more anterior portions of the middle temporal gyri, STS, and STG during processing of speech, while symbolic gestures activated the fusiform gyri, ventral inferior temporal, and inferior occipital gyri. Similarly, connectivity analyses showed that the left IFG and pMTG were coupled to the superior temporal gyri only for speech, to fusiform and inferior temporal gyri only for symbolic gesture. These anterior, superior (speech-related), and inferior temporal (gesture-related) areas, which lie on either side of the modality-independent conjunctions, are in closer proximity to unimodal auditory and visual cortices respectively.

We suggest that these anterior and inferior temporal areas may be involved in extracting higher level but still modality-dependent information encoded in speech or symbolic gesture. These regions may process domain-specific semantic features (that were not reproduced in our control conditions) including phonological or complex acoustic features of voice or speech on the one hand (e.g., ref. 34) and complex visual attributes of manual or facial stimuli on the other (e.g., ref. 35).

Synthesis: A Model for Processing of Symbolic Gesture and Spoken Language. Taken together, the conjunction, task contrast, and connectivity data support a model in which modality-specific areas in the superior and inferior temporal lobe extract salient features from vocal-auditory or gestural-visual signals respectively. This information then gains access to a modality-independent perisylvian system where it is mapped onto the corresponding conceptual representations. The posterior temporal regions provide contact between symbolic gestures or spoken words and the semantic features they encode, while inferior frontal regions guide a rule-based process of selection and integration with aspects of world knowledge that may be more widely distributed throughout the brain. In this model, inferior frontal and posterior temporal areas correspond to an amodal system (36) that plays a central role in human communication—a semiotic network in which meaning (the signified) is paired with symbols (the signs) whether these are words, gestures, images, sounds, or objects (37). This is not a novel concept. Indeed, Broca (38) proposed that language was part of a more comprehensive communication system that supported a broader “ability to make a firm connection between an idea and a symbol, whether this symbol was a sound, a gesture, an illustration or any other expression.”

Caveats. Before discussing the possible implications of these results, a number of caveats need to be considered. First, our results relate only to comprehension. Shared and unique features might differ considerably for symbolic gesture and speech production. Similarly, the present findings pertain only to overlaps between symbolic gesture and language at the corresponding level of words, phrases, or simple argument structures. The neural architecture may differ dramatically at more complex linguistic levels, and regions that are recruited in recursive or compositional processes may never be engaged during the processing of pantomimes or emblems (see *SI Note 5*).

Although the autonomous gestures we presented are dissociable from language—and were indeed selected for this reason—it is possible that perisylvian areas were activated simply because the meanings encoded in the gestures were being recoded into language. This possibility cannot be ruled out, although we consider it unlikely. Historically, similar concerns may have been raised about investigations of sign language comprehension in hearing participants (6); however, these results are commonly considered to be real rather than artifacts of recoding or translation. Moreover, in the present experiment, we used normative data to select only emblems or pantomimes that were very familiar to English speakers (see *SI Methods*) and that were already part of their communicative repertoire and should not have led to verbal labeling as novel or unfamiliar gesture might have been. Additionally, to ensure that recoding did not occur, we chose not to superimpose any task that would have encouraged this, such as requiring participants to remember the gestures or make a decision about them, i.e., tasks that might have been more effectively executed by translation of gestures into spoken language.

Finally, the results of our analyses themselves argue against recoding: if conjunctions in the inferior frontal and posterior temporal regions reflect only verbal labeling, then the gesture-specific (gesture > speech) activations would be expected to encompass the entire network of brain regions that process symbolic gestures. This contrast highlighted only a relatively small portion of the inferior temporal and fusiform gyri, which would not likely constitute the complete the symbolic gesture-processing network, should one exist. Nor did the gesture-specific activations include regions that have been shown to be activated during language translation or switching (39), areas that would likely have been engaged during the process of recoding into English. Nevertheless, to address this issue directly, it would be interesting to perform follow-up studies in which the likelihood of verbal recoding could be manipulated parametrically (see *SI Note 6*).

Implications. Respecting the caveats listed above, our findings have important implications for theories of the use, development, and potentially the evolutionary origins of human communication. The fact that gestural and linguistic milestones are consistently and closely linked during early development, tightly coupled in both deaf and hearing children (e.g., ref. 40) might be best understood if they were instantiated in a common neural substrate that supports symbolic communication (see *SI Note 7*).

The concept of a common substrate for symbolic gesture and language may bear upon the question of language evolution of as well, providing a possible mechanism for the gestural origins hypothesis, which proposes that spoken language emerged through adaptation of a gestural communication system that existed in a common ancestor (see *SI Note 8*).

A popular current theory offers an anatomical substrate for this hypothesis, proposing that the mirror neuron system, activated during observation and execution of goal directed actions (41), constituted the language-ready system in the early human brain. This provocative idea has gained widespread appeal in past several years but is not universally accepted and may not adequately account for the more complex semantic properties of natural language (42, 43).

However, the mirror neuron system does not constitute the only plausible language-ready system consistent with a gestural origins account. The present results provide an alternative (or complementary) possibility. Rather than being fundamentally related to surface behaviors such as grasping or ingestion or to an automatic

sensorimotor resonance between physical actions and meaning, language may have developed (at least in part) upon a different neural scaffold, i.e., the same perisylvian system that presently processes language. In this view, a precursor of this system supported gestural communication in a common ancestor, where it played a role in pairing gesture and meaning. It was then adapted for the comparable pairing of sound and meaning as voluntary control over the vocal apparatus was established and spoken language evolved. And—as might be predicted by this account—the system continues to process both symbolic gesture and spoken language in the human brain.

Materials and Methods

Blood oxygenation level-dependent contrast images were acquired in 20 healthy right-handed native English speakers as they observed video clips of an actress performing gestures (60 emblems, 60 pantomimes) and the same number of clips in which the actress produced the corresponding spoken glosses (expressing the gestures' meaning in English). Control stimuli, designed to reproduce corresponding surface features—from lower level visual and auditory elements to higher level attributes such as human form and biological motion—were also presented. Task effects were estimated using a general linear model in which expected task-related responses were convolved with a standard hemodynamic response function. Random effects conjunction and contrast analyses as well as functional connectivity analyses using seed region linear regressions were performed. (Detailed descriptions are provided in *SI Methods*).

ACKNOWLEDGMENTS. The authors thank Drs. Susan Goldin-Meadow and Jeffrey Solomon for their important contributions. This research was supported by the Intramural program of the NIDCD.

- Suppalla T (2003) Revisiting visual analogy in ASL classifier predicates. *Perspectives on classifier constructions in sign language*, ed Emmorey K (Lawrence Erlbaum Assoc, Mahwah, NJ), pp 249–527.
- Petitto LA (1987) On the autonomy of language and gesture: Evidence from the acquisition of personal pronouns in American Sign Language. *Cognition* 27:1–52.
- McNeill D (1992) *Hand and Mind* (University of Chicago Press, Chicago), pp 36–72.
- Kendon A (1988) How gestures can become like words. *Crosscultural Perspectives in Nonverbal Communication*, ed Poyatos F (Hogrefe, Toronto), pp 131–141.
- Studdert-Kennedy M (1987) The phoneme as a perceptuomotor structure. *Language Perception and Production*, eds Allport A, MacKay D, Prinz W, Scheerer E (Academic, London), pp 67–84.
- MacSweeney M, Capek CM, Campbell R, Woll B (2008) The signing brain: The neurobiology of sign language. *Trends Cogn Sci* 12:432–440.
- Rumiati RI, et al. (2004) Neural basis of pantomiming the use of visually presented objects. *Neuroimage* 21:1224–1231.
- Villarreal M, et al. (2008) The neural substrate of gesture recognition. *Neuropsychologia* 46:2371–2382.
- Gallagher HL, Frith CD (2004) Dissociable neural pathways for the perception and recognition of expressive and instrumental gestures. *Neuropsychologia* 42:1725–1736.
- Knutson KM, McClellan EM, Grafman J (2008) Observing social gestures: An fMRI study. *Exp Brain Res* 188:187–198.
- Montgomery KJ, Haxby JV (2008) Mirror neuron system differentially activated by facial expressions and social hand gestures: A functional magnetic resonance imaging study. *J Cogn Neurosci* 20:1866–1877.
- Bernardis P, Gentilucci M (2006) Speech and gesture share the same communication system. *Neuropsychologia* 44:178–190.
- Demonet JF, Thierry G, Cardebat D (2005) Renewal of the neurophysiology of language: functional neuroimaging. *Physiol Rev* 85:49–95.
- Jobard G, Vigneau M, Mazoyer B, Tzourio-Mazoyer N (2007) Impact of modality and linguistic complexity during reading and listening tasks. *Neuroimage* 34:784–800.
- Decety J, et al. (1997) Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain* 120:1763–1777.
- Emmorey K, Xu J, Gannon P, Goldin-Meadow S, Braun AR (2010) CNS activation and regional connectivity during pantomime observation: No engagement of the mirror neuron system for deaf signers. *Neuroimage* 49:994–1005.
- Lotze M, et al. (2006) Differential cerebral activation during observation of expressive gestures and motor acts. *Neuropsychologia* 44:1787–1795.
- Damasio AR (1992) Aphasia. *N Engl J Med* 326:531–539.
- Bookheimer S (2002) Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annu Rev Neurosci* 25:151–188.
- Koenig P, et al. (2005) The neural basis for novel semantic categorization. *Neuroimage* 24:369–383.
- Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ (1997) Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proc Natl Acad Sci USA* 94:14792–14797.
- Hagoort P (2005) On Broca, brain, and binding: A new framework. *Trends Cogn Sci* 9:416–423.
- Martin A (2007) The representation of object concepts in the brain. *Annual Review of Psychology* 58:25–45.
- Gitelman DR, Nobre AC, Sonty S, Parrish TB, Mesulam MM (2005) Language network specializations: An analysis with parallel task designs and functional magnetic resonance imaging. *Neuroimage* 26:975–985.
- Indefrey P, Levelt WJ (2004) The spatial and temporal signatures of word production components. *Cognition* 92:101–144.
- Hart J, Jr, Gordon B (1990) Delineation of single-word semantic comprehension deficits in aphasia, with anatomical correlation. *Ann Neurol* 27:226–231.
- Saygin AP, Dick F, Wilson SM, Dronkers NF, Bates E (2003) Neural resources for processing language and environmental sounds: Evidence from aphasia. *Brain* 126:928–945.
- Beauchamp MS, Lee KE, Argall BD, Martin A (2004) Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron* 41:809–823.
- Okada K, Hickok G (2009) Two cortical mechanisms support the integration of visual and auditory speech: A hypothesis and preliminary data. *Neurosci Lett* 452:219–223.
- Skipper JJ, Nusbaum HC, Small SL (2005) Listening to talking faces: Motor cortical activation during speech perception. *Neuroimage* 25:76–89.
- Lau EF, Phillips C, Poeppel D (2008) A cortical network for semantics: (De)constructing the N400. *Nat Rev Neurosci* 9:920–933.
- Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8:393–402.
- Gunter TC, Bach P (2004) Communicating hands: ERPs elicited by meaningful symbolic hand postures. *Neurosci Lett* 372:52–56.
- Rogalsky C, Hickok G (2009) Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cereb Cortex* 19:786–796.
- Corina D, et al. (2007) Neural correlates of human action observation in hearing and deaf subjects. *Brain Res* 1152:111–129.
- Riddoch MJ, Humphreys GW, Coltheart M, Funnel E (1988) Semantic system or systems? Neuropsychological evidence re-examined. *Cognit Neuropsychol* 5:3–25.
- Eco U (1976) *A Theory of Semiotics* (Indiana Univ Press, Bloomington).
- Broca P (1861) Remarques sur le siege de la faculte du langage articule; suivies d'une observation d'aphemie. *Bulletin de la Societe Anatomique de Paris* 6:330–357.
- Price CJ, Green DW, von Studnitz R (1999) A functional imaging study of translation and language switching. *Brain* 122:2221–2235.
- Bates E, Dick F (2002) Language, gesture, and the developing brain. *Dev Psychobiol* 40:293–310.
- Rizzolatti G, Arbib MA (1998) Language within our grasp. *Trends Neurosci* 21:188–194.
- Hickok G (2009) Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J Cogn Neurosci* 21:1229–1243.
- Toni I, de Lange FP, Noordzij ML, Hagoort P (2008) Language beyond action. *Journal of Physiology-Paris* 102:71–79.