

# Converging Language Streams in the Human Temporal Lobe

Galina Spitsyna,<sup>1\*</sup> Jane E. Warren,<sup>2\*</sup> Sophie K. Scott,<sup>3</sup> Federico E. Turkheimer,<sup>2</sup> and Richard J. S. Wise<sup>1,2</sup>

<sup>1</sup>University Department of Clinical Neurosciences, Royal Free and University College Medical School, London NW3 2PF, United Kingdom, <sup>2</sup>Division of Neuroscience and Mental Health and Medical Research Council Clinical Sciences Centre, Imperial College London, Hammersmith Hospital, London W12 0NN, United Kingdom, and <sup>3</sup>Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, United Kingdom

There is general agreement that, after initial processing in unimodal sensory cortex, the processing pathways for spoken and written language converge to access verbal meaning. However, the existing literature provides conflicting accounts of the cortical location of this convergence. Most aphasic stroke studies localize verbal comprehension to posterior temporal and inferior parietal cortex (Wernicke's area), whereas evidence from focal cortical neurodegenerative syndromes instead implicates anterior temporal cortex. Previous functional imaging studies in normal subjects have failed to reconcile these opposing positions. Using a functional imaging paradigm in normal subjects that used spoken and written narratives and multiple baselines, we demonstrated common activation during implicit comprehension of spoken and written language in inferior and lateral regions of the left anterior temporal cortex and at the junction of temporal, occipital, and parietal cortex. These results indicate that verbal comprehension uses unimodal processing streams that converge in both anterior and posterior heteromodal cortical regions in the left temporal lobe.

**Key words:** speech; reading; PET; temporal pole; fusiform gyrus; temporoparietal cortex

## Introduction

Although the perception of spoken and written language depends on separate modality-specific sensory systems, psychological models postulate convergence on common systems for decoding the meaning of words, phrases, and sentences. The premise that cognitive processing is dependent on anatomical convergence of sensory pathways has been most clearly expressed in a model that describes the processing of sensory experience in terms of a cortical hierarchy: from primary sensory cortex, through unimodal and then heteromodal association cortices, and thence to paralimbic and limbic regions (Mesulam, 1998). This model, which owes much to the proposals of Geschwind (1965), emphasizes the role of heteromodal association cortex in processing convergent sensory information. Metabolic brain imaging in the nonhuman primate has demonstrated convergence of auditory and visual pathways in inferior parietal cortex, the superior temporal sulcus, medial temporal cortex, and prefrontal cortex (Poremba et al., 2003). The wide distribution of these heteromodal regions reflects the differing projections of functionally distinct dorsal and ventral pathways within visual and auditory cortex (Ungerleider and Mishkin, 1982; Rauschecker and Tian, 2000). Analogous auditory and visual pathways have been demonstrated in the human brain (Ungerleider and Haxby, 1994; Alain et al., 2001; Scott and Johnsrude, 2003).

The stroke literature emphasizes the role of posterior hetero-

modal cortex in language comprehension (Geschwind, 1965; Kertesz et al., 1982; Alexander et al., 1989). A number of human functional imaging studies support this position, by demonstrating activation in posterior temporal or inferior parietal cortex in response to lexical or semantic tasks (for review, see Binder et al., 2003). However, other studies have implicated inferior and lateral regions of anterior temporal cortex (Scott et al., 2000, 2003; Marinkovic et al., 2003; Sharp et al., 2004). Clinical studies of the neurodegenerative disorder semantic dementia, which is characterized by cortical atrophy predominantly affecting the anterior and ventral temporal lobe, have demonstrated a progressive impairment of language comprehension in the context of impaired semantic knowledge across nonverbal domains (Hodges et al., 1992; Bozeat et al., 2000; Chan et al., 2001; Gorno-Tempini et al., 2004). Therefore, different studies have implicated either the posterior or anterior temporal cortex in language comprehension. These separate opinions, polarized along the posterior–anterior axis of the temporal lobe, have been expressed in recent influential review articles (McClelland and Rogers, 2003; Hickok and Poeppel, 2004; Catani and ffytche, 2005).

We propose that access to verbal meaning does not occur via either anterior or posterior temporal cortex alone, but depends on both anterior and posterior heteromodal cortical systems within the temporal lobe. To test this hypothesis, we performed a functional imaging study that investigated the anatomical convergence of spoken and written language processing in the human brain. We chose to use positron emission tomography (PET), because it more reliably records signal in ventral temporal cortex than functional magnetic resonance imaging using echo-planar imaging (Devlin et al., 2000). In addition, we used multiple baseline conditions to optimize interpretation of the functional imaging data (Friston et al., 1996; Binder et al., 1999).

Received Feb. 7, 2006; revised May 2, 2006; accepted May 29, 2006.

This work was supported by the Wellcome Trust, United Kingdom (G.S.), and by Action Medical Research, United Kingdom, and the Barnwood House Trust, United Kingdom (J.E.W.).

\*G.S. and J.E.W. contributed equally to this work.

Correspondence should be addressed to Prof. Richard J. S. Wise, Medical Research Council Clinical Sciences Centre, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK. E-mail: richard.wise@csc.mrc.ac.uk.

DOI:10.1523/JNEUROSCI.0559-06.2006

Copyright © 2006 Society for Neuroscience 0270-6474/06/267328-09\$15.00/0

## Materials and Methods

**Participants.** Eleven healthy right-handed unpaid volunteers (three females; age range, 37–82 years; mean, 52 years), with normal or corrected-to-normal vision and normal hearing for age, participated in the study. The studies were approved both by the Administration of Radioactive Substances Advisory Committee (Department of Health, United Kingdom) and by the local research ethics committee of the Hammersmith Hospitals NHS Trust.

**Behavioral tasks.** Intelligible spoken and written narratives were contrasted with unintelligible auditory (spectrally rotated speech) and visual (false font) stimuli matched closely to the narratives in terms of auditory and visual complexity. Thus, during data analyses, the contrasts of the intelligible with the modality-specific unintelligible conditions were intended to demonstrate processing of speech and written language beyond the cortical areas specialized for early auditory and visual processing. Spoken and written narratives were extracts taken from a popular nonfiction book aimed at a general adult readership. Digital recordings of spoken passages (each ~1 min long) were made in an anechoic chamber using a single female speaker, with pauses, stress, and pitch appropriate to the meaning of each phrase and sentence. Written narratives (each ~200 words long) consisted of a single block of paragraph-formatted text. The modality-specific unintelligible auditory baseline stimuli were created by spectral rotation of the digitized spoken narratives (Scott et al., 2000), using a digital version of the simple modulation technique described by Blesser (1972). Spectrally rotated speech preserves the acoustic complexity, the intonation patterns, and some of the phonetic cues and features of normal speech; however, it is not intelligible without many hours of training, and speaker identity cues are largely abolished. The unintelligible visual stimuli were created using “false font,” consisting of symbols from the International Phonetic Alphabet. The false font was grouped to form text-like arrays, including punctuation marks, to match the written texts in overall visual appearance. Subjects were presented with short examples of rotated speech and false font immediately before the scanning session, but no subject had experience with either rotated speech or the phonetic alphabet before participation in this study. During scanning, stimuli were presented on an Apple laptop computer using Pyscope software. For speech and rotated speech, presented via insert earphones, subjects were instructed to listen attentively. For written text and false font, presented on a computer screen in black size 12 Times font on a white background, subjects were instructed to read the texts silently and to scan from left to right across the lines of false font as if they were reading. Each intelligible passage, spoken or written, was presented only once.

It has been suggested previously that subjects whose attention is not held by perception of meaningful stimuli or by performance of an explicit task may generate stimulus-independent thoughts or recollections, thereby activating brain regions involved in retrieval of declarative memory, including semantic knowledge (Binder et al., 1999; Gusnard and Raichle, 2001; Mazoyer et al., 2001; Stark and Squire, 2001). We hypothesized that subjects might experience stimulus-independent thoughts during passive perception of the unintelligible baseline stimuli, resulting in some masking of activation associated with semantic processing of intelligible narratives during direct comparisons of narrative and baseline conditions. To optimize detection of activation associated with verbal semantic processing, we used an attention-demanding numerical task as an additional baseline. Use of this baseline task has previously been shown to unmask inferomedial temporal lobe activity associated with processing of meaningful stimuli (Stark and Squire, 2001). Functional imaging evidence demonstrates that conceptual knowledge about numbers is not stored in the temporal lobes, but in parietal regions centered on the inferior parietal sulcus (Thioux et al., 2005), so that a number task would not be expected to activate temporal lobe regions involved in semantic aspects of language processing. During the number task, the subjects heard a series of numbers randomly selected from the integers 1–10, and were instructed to press one of two buttons with the left hand after each number to indicate whether the number was odd or even. During this task, the rate of stimulus presentation varied according to the subject's speed of response, because each button press resulted in

the immediate presentation of the next number; therefore, subjects had no rest time during this task. Equal emphasis was placed on response speed and accuracy. Subjects performed with >85% accuracy. Because the number task was conducted only in the auditory modality, and the number of words presented per unit time was less in this than in the other auditory conditions, contrasts of the language stimuli with this condition were not expected to subtract out activity associated with early auditory and visual processing of intelligible language. Rather, contrasts with the number task were used to reveal upstream high-order association cortex, insensitive to input modality, that processed meaning conveyed by intelligible language.

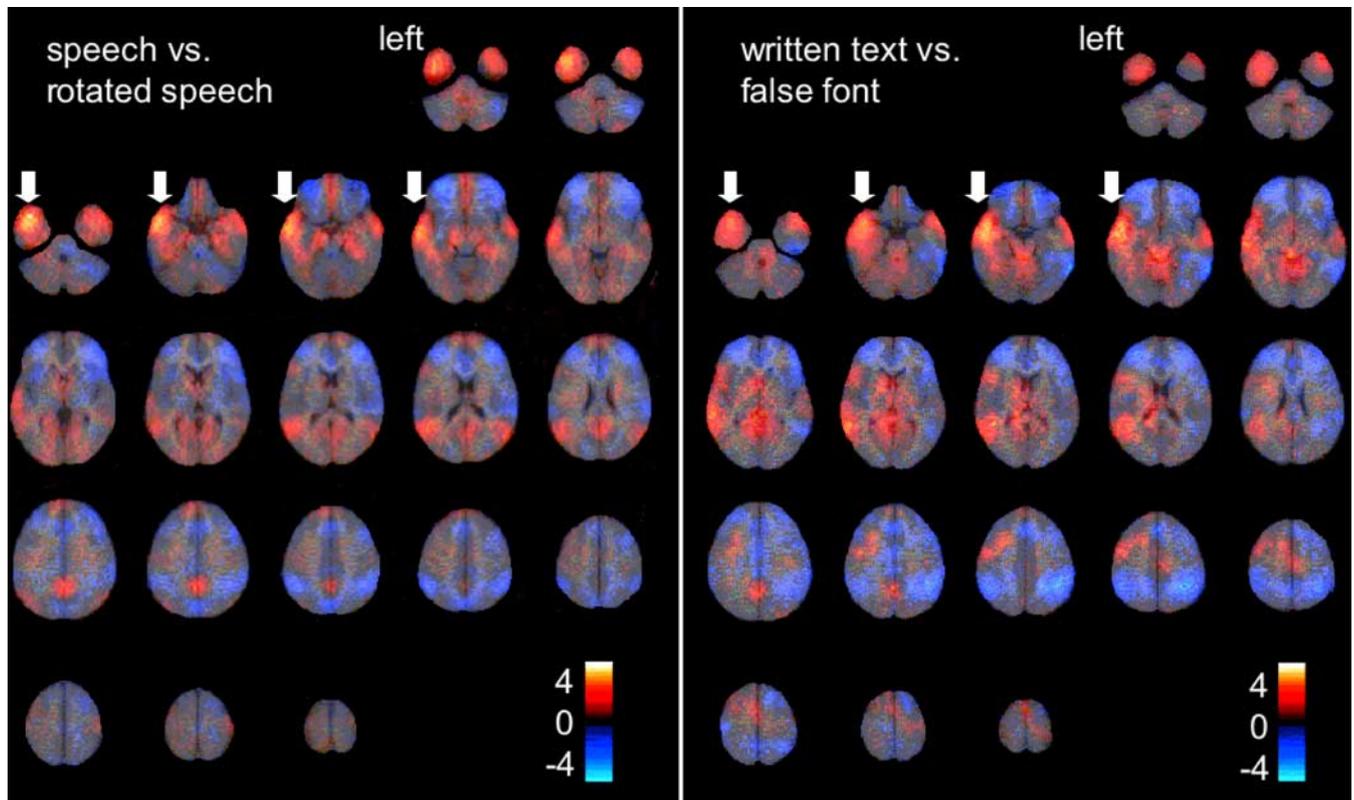
**Functional neuroimaging.** The subjects were scanned on a Siemens (Erlangen, Germany) HR++ (966) PET camera operated in high-sensitivity three-dimensional mode. The field of view (20 cm) covers the whole brain with a resolution of 5.1 mm full width at half-maximum in *x*-, *y*-, and *z*-axes. A transmission scan was performed for attenuation correction. Sixteen scans were performed on each subject, with the subjects' eyes closed during the auditory conditions. Three scans each were allocated to the listening and reading conditions and their auditory and visual baseline conditions, and four scans were allocated to the number task. Condition order was randomized within and between subjects.

The dependent variable in functional imaging studies is the hemodynamic response: a local increase in synaptic activity is associated with increased local metabolism, coupled to an increase in regional cerebral blood flow (rCBF). Water labeled with a positron-emitting isotope of oxygen ( $H_2^{15}O$ ) was used as the tracer to demonstrate changes in rCBF, equivalent to changes in tissue concentration of  $H_2^{15}O$ . During each scan, ~5 mCi of  $H_2^{15}O$  was infused as a slow bolus over 40 s, resulting in a rapid rise in measurable emitted radioactivity (head counts) that peaks after 30–40 s. Stimulus presentation encompassed the incremental phase, by commencing 10 s before the rise in head counts and continuing for 10–15 s after the counts began to decline as a result of washout and radioactive decay. Individual scans were separated by intervals of 6 min.

**Statistical modeling in wavelet space.** The contrast of speech with its auditory baseline (rotated speech) and the contrast of reading with its visual baseline (false font) were statistically modeled in wavelet space (Turkheimer et al., 2000; Aston et al., 2005). The resulting images demonstrated the distribution of estimated effect size (expressed as *Z*-scores) across the brain, both positive (intelligible language > perceptual baseline) and negative (perceptual baseline > intelligible language), without a cutoff threshold.

**Whole-brain subtractive analyses.** Whole-brain statistical analyses were performed using SPM99 software (Wellcome Department of Cognitive Neurology, London, UK; <http://www.fil.ion.ucl.ac.uk/spm>). We used a fixed-effects model to generate statistical parametric maps representing the results of voxel-wise *t* test comparisons for contrasts of interest. The voxel-level statistical threshold was set at  $p < 0.05$ , with family-wise error (FWE) correction for multiple comparisons, and a cluster extent threshold of 10 voxels. Inclusive masking was used to identify activation regions common to more than one contrast, with a statistical threshold equivalent to  $p < 0.05$  (*Z*-score, >4.8), FWE-corrected, for each contrast (Nichols et al., 2005). Use of a fixed-effects model allowed relative activation for all conditions to be directly visualized and evaluated, via effect size histograms, in all contrasts of interest. Describing regional activity during one condition relative to activity across all conditions avoids some of the potential errors of interpretation inherent in cognitive subtractions (Friston et al., 1996).

**Region of interest analyses.** For the region of interest (ROI) analyses, we defined three regions in the left hemisphere and two in the right hemisphere. Informed by our previous studies of the implicit and explicit processing of spoken and written language (Scott et al., 2000, 2003; Crinion et al., 2003; Sharp et al., 2004), we selected anatomical ROIs for the left lateral temporal pole (TP) and anterior fusiform gyrus (FG), and their homotopic regions in the right hemisphere, from a probabilistic electronic brain atlas (Hammers et al., 2003). In the absence of probabilistic anatomical data for a posterior region at the junction of the temporal, occipital, and parietal lobes (the TOP junction), identified by Binder et al. (2003) as a region involved in access to single word meaning during the performance of explicit lexical and semantic tasks, we defined this



**Figure 1.** The results of analyses within wavelet space of two contrasts: speech with rotated speech, and written text with false font. The estimated effect sizes, both positive (red–yellow) and negative (blue) results, expressed as Z-scores without a cutoff threshold, are displayed on axial brain slices from ventral to dorsal. The strongest positive effect size for each contrast was observed within the left anterior temporal cortex (arrows).

ROI on functional grounds as the voxels activated in common in the left hemisphere in the contrasts of spoken and written language with the number task. An ROI for the right TOP junction was not defined, because there was no objective method of identifying a homotopic region in the right hemisphere. The ROI analyses were performed using the MarsBaR software toolbox within SPM99 (Brett et al., 2002). Individual measures of mean effect size for both language conditions and their modality-specific baselines (all contrasted with the common baseline of the number task) were obtained for each region in each subject. These measures were entered into repeated-measures ANOVAs. Within-region two-way ANOVAs used the factors “modality” and “intelligibility”; between-region three-way ANOVAs used the additional factor of “region” or “hemisphere,” as appropriate. Because the number of ANOVAs performed was  $>5$  and  $<10$ , significance was set at  $p < 0.005$ . Significant effects or interactions were further examined using *post hoc* paired *t* test comparisons, with significance set at  $p < 0.05$ .

## Results

### Initial overview

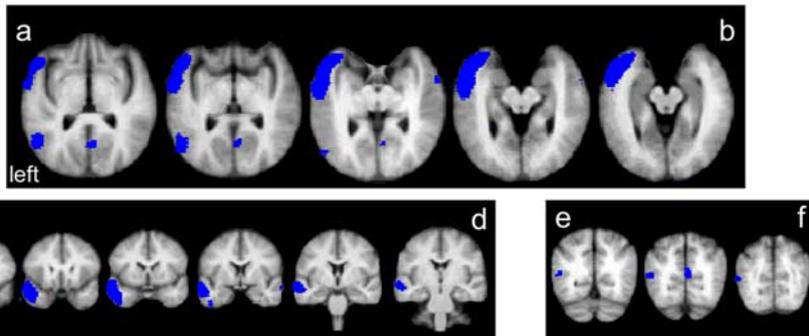
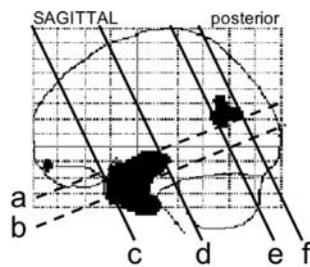
The analysis in wavelet space gave an overview of activity associated with the two modality-specific contrasts: speech versus rotated speech, and reading versus false font (Fig. 1). The effect size maps gave an overall impression of the regional effects associated with the implicit processing of auditory and visual language. This preliminary analysis guarded against accepting apparent functional dissociations between regions that can attract the eye when viewing thresholded statistical parametric maps (Jernigan et al., 2003). Two main conclusions were evident from this analysis. The first was that, for both contrasts, the largest positive estimated effect size was observed in the left anterior and ventral temporal lobe. However, there was also activity evident in homotopic cortex on the right, and the demonstration of a functional

hemisphere asymmetry in these regions was dependent on a formal statistical comparison between left and right anterior and inferior temporal ROIs (Jernigan et al., 2003). The second conclusion was that, other than in the anterior temporal lobes, activations were more clearly left-lateralized for reading than listening to speech.

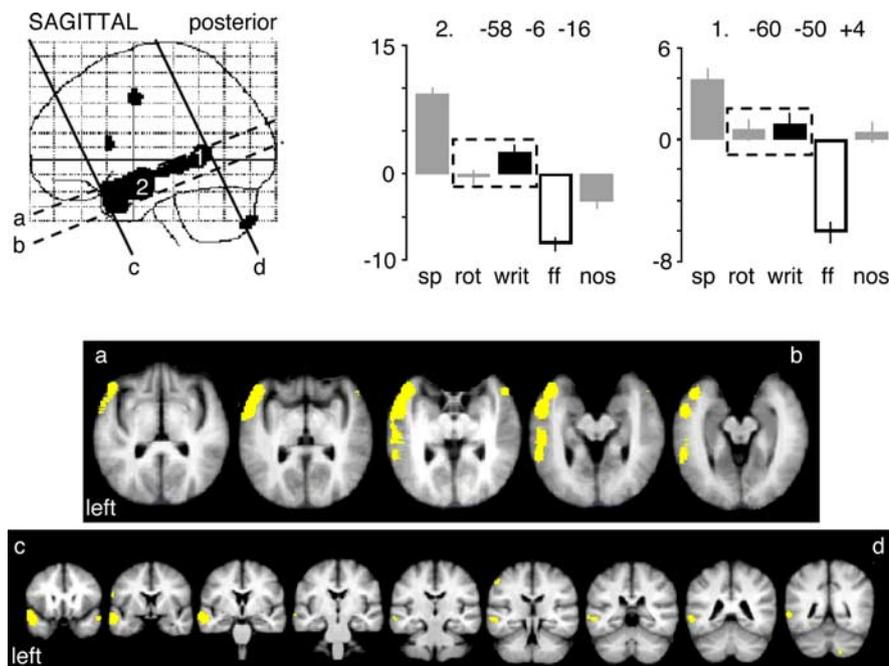
### Within-modality responses to intelligible language

The explicit aim of this study was to investigate temporal lobe regions engaged in the processing of intelligible spoken and written language. Although the contrasts described below also demonstrated activations outside the temporal cortices, these will be mentioned but not discussed further. Using SPM99, we investigated regions engaged in the processing of spoken language by contrasting speech with the rotated speech baseline. This contrast, identified preferential responses to intelligible speech within anterior and posterior left temporal neocortex (Fig. 2). The larger was an anterolateral temporal region that included the anterior superior temporal sulcus (STS) and the lateral TP and extended ventrally into inferotemporal cortex. Additional activation was observed in posterior STS. Speech-related activity in left mid-STS did not survive the statistical threshold set for the subtractive analyses (voxel-level Z-scores,  $>3.4$  but  $<4.8$ ).

We then investigated regions engaged in the processing of written language by contrasting written text with false font. This demonstrated a response along the length of the left STS (Fig. 3), extending into the lateral TP. The profiles of responses across all conditions demonstrated an effect of modality as well as intelligibility, with no difference (Z-score,  $<1.5$ ) between the responses



**Figure 2.** The contrast of speech with rotated speech, displayed on a sagittal statistical parametric map (SPM), and coregistered (as a blue overlay) on to brain slices from a mean magnetic resonance imaging (MRI) template in Montreal Neurological Institute stereotactic space. The MRI slices were oriented along the plane of the STS, and consecutive axial (a–b) and coronal (c–d and e–f) slices, separated by 4 mm, are displayed with reference to their location on the SPM.



**Figure 3.** Written text contrasted with false font, using the same display design as in Figure 2 and a yellow overlay on the coregistered slices. The response profiles across all five conditions are shown for two peak voxels, identified within SPM99, at either end of the left STS. The size of effect (in arbitrary units) for the conditions used in the contrast, written text (writ; black bar) and false font (ff; white bar), are shown in relation to activity in response to speech (sp), rotated speech (rot), and the number task (nos), shown as gray bars. Mean activity ( $\pm$  SEM) for all five conditions was normalized around zero. Coordinates (in millimeters) show the location of peaks in Montreal Neurological Institute space, located within the posterior STS (1) and anterior STS (2). The dashed outlines draw attention to the equivalent response to rotated speech and written text at the two peak voxels.

to unintelligible rotated speech and to intelligible written text in voxels within the left STS posterior to the TP.

Because the direct contrast of written text with false font alone did not identify activity in a ventral visual stream of processing

any of the three regions. The apparent functional association between the anatomically widely separated anterior FG and TOP junction was formally investigated by a pairwise comparison, using a 2 (modality)  $\times$  2 (intelligibility)  $\times$  2 (region)

for written words within the temporal lobes, as had been demonstrated in the electrophysiological study of Nobre et al. (1994), an additional contrast of written text with both the unintelligible conditions, auditory and visual, was performed (Fig. 4). This identified activation along the length of the left FG, extending into both inferior and lateral regions of anterior temporal cortex. The response profiles at activation peaks along the left FG showed that activity in response to false font declined progressively the more anterior the location of the activation peak (Fig. 4).

**Across-modality responses to intelligible language**

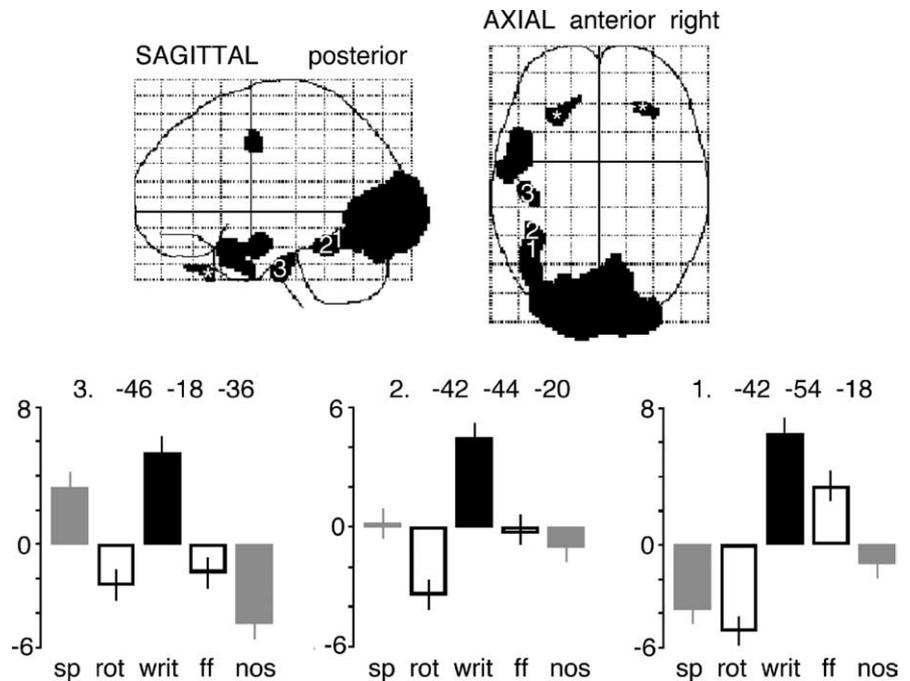
We identified voxels that were common to the contrasts of speech with rotated speech and written text with false font (Fig. 5). As expected from the individual contrasts, a large region of common activation was observed in the left anterior temporal cortex, located predominantly within the lateral TP. Because the perception of unintelligible stimuli without task demand may be associated with stimulus-independent processes, including both episodic and semantic memory, we examined the contrasts of spoken and written language with the number task (Stark and Squire, 2001). In addition to activity within the left lateral TP that had been identified in the previous comparison, activation common to processing spoken and written language was identified in the left anterior FG, and at the junction of the TOP junction, just ventral to the angular gyrus (Fig. 5). There were additional common activations in midline cortex: posterior to the splenium of the corpus callosum, in the superior frontal gyrus and in medial orbitofrontal cortex (Fig. 5).

Mean activity across subjects in the three left hemisphere ROIs were entered into separate 2 (modality)  $\times$  2 (intelligibility) ANOVAs. Within the lateral TP, there were significant main effects of both modality (auditory > visual,  $F_{(1,10)} = 27.9$ ,  $p < 0.0005$ ) and intelligibility (intelligible > unintelligible,  $F_{(1,10)} = 40.6$ ,  $p < 0.0005$ ) (Fig. 6). Within the anterior FG and the TOP junction, there was a main task effect of intelligibility alone (intelligible > unintelligible,  $F_{(1,10)} = 21.2$ ,  $p < 0.001$ , and  $F_{(1,10)} = 20.7$ ,  $p < 0.001$ , respectively) (Figs. 6, 7). There was no significant modality by intelligibility interaction within

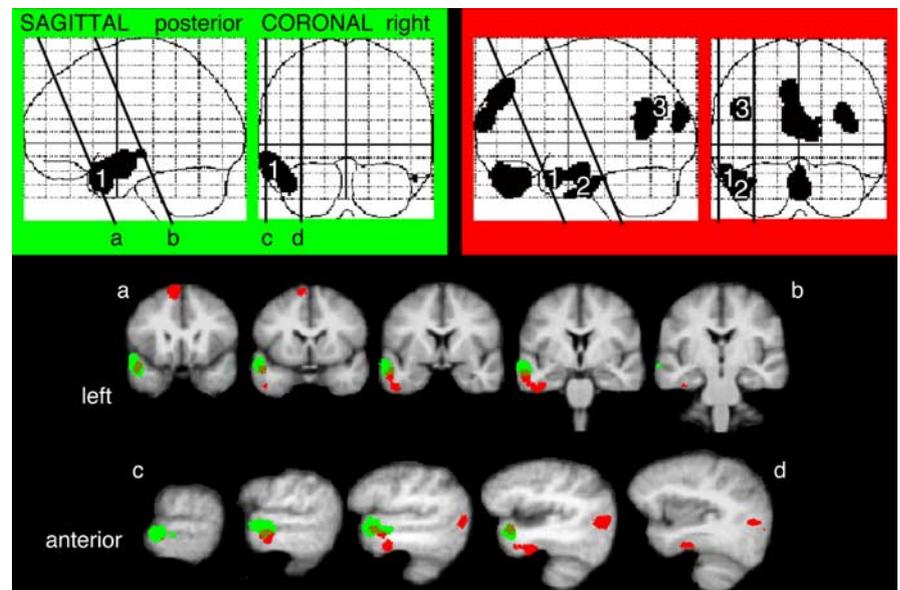
ANOVA: this confirmed that there was no region by intelligibility interaction ( $F_{(1,10)} = 1.8$ ;  $p > 0.2$ ) nor region by modality interaction ( $F_{(1,10)} = 3.6$ ;  $p = 0.09$ ).

The data from the left hemisphere ROIs demonstrated the response to the modality-specific baseline conditions within these three regions. Based on the 95% confidence intervals, the response to both modality-specific baseline conditions, relative to the number task, was significant within the left TOP junction (Fig. 7), with a strong trend toward significance within the left anterior FG (Fig. 6). In contrast, there was no activity in response to the modality-specific baseline conditions in the left lateral TP (in fact, the response to false font was significantly less than to the number task, which can be attributed, at least in part, to the effect of modality, because the number task used spoken numbers) (Fig. 6).

The ROI analyses were also used to investigate hemispheric asymmetries between left and right lateral TP and anterior FG, using separate 2 (modality)  $\times$  2 (intelligibility)  $\times$  2 (hemisphere) ANOVAs. Between the left and right lateral TP, there was no main effect of hemisphere but a strong trend toward a significant intelligibility by hemisphere interaction ( $F_{(1,10)} = 13.6$ ;  $p < 0.01$ ). *Post hoc* paired *t* tests on the differences in response between each intelligible language condition and its modality-specific baseline showed a significantly greater effect within the left compared with the right lateral TP for both the auditory and visual modalities:  $t_{(10)} = 3.9$ ,  $p < 0.005$ ; and  $t_{(10)} = 2.7$ ,  $p < 0.05$ , respectively. Although the response to speech was the same in left and right lateral TP, the response to rotated speech was asymmetrical (right > left, *post hoc* paired *t* test,  $t_{(10)} = 2.6$ ;  $p < 0.05$ ) (Fig. 6); indicating that there are auditory cues and features common to speech and rotated speech that activate the right lateral TP more strongly than the left. In the anterior FG, there was no significant main effect of hemisphere. However, there was a trend toward significance for an intelligibility by hemisphere interaction ( $F_{(1,10)} = 6.4$ ;  $p < 0.05$ ), and a significant modality by hemisphere interaction ( $F_{(1,10)} = 37.0$ ;  $p < 0.0005$ ), the result of a symmetrical response to speech, but an asymmetrical response to reading (left > right, *post hoc* paired *t* test:  $t_{(10)} = 4.1$ ;  $p < 0.005$ ). Despite the absence of a probabilistic atlas to objectively define homotopic cortex in the right hemisphere to pair with the left TOP junction, nevertheless the separate whole-brain analyses for speech and reading contrasted with the number task did indicate a symmetrical response to speech and an asymmetrical (left > right) response to reading (Fig. 7). These analyses confirm a preferential, but not



**Figure 4.** A display of the statistical parametric map, both sagittal and axial views, for the contrast of intelligible written text (writ) with the two unintelligible baseline conditions, rotated speech (rot) and false font (ff). Activity across all five conditions is taken from peak voxels along the length of the left FG, using the same style of display as in Figure 3. At voxel 1, with stereotactic coordinates cited as the location of the visual word form area (Cohen and Dehaene, 2004), there was still a response to false font relative to rotated and intelligible speech, but this was considerably reduced or absent at voxels 2 and 3. The regions marked with asterisks show activity within extraocular muscles, active during the visual conditions.

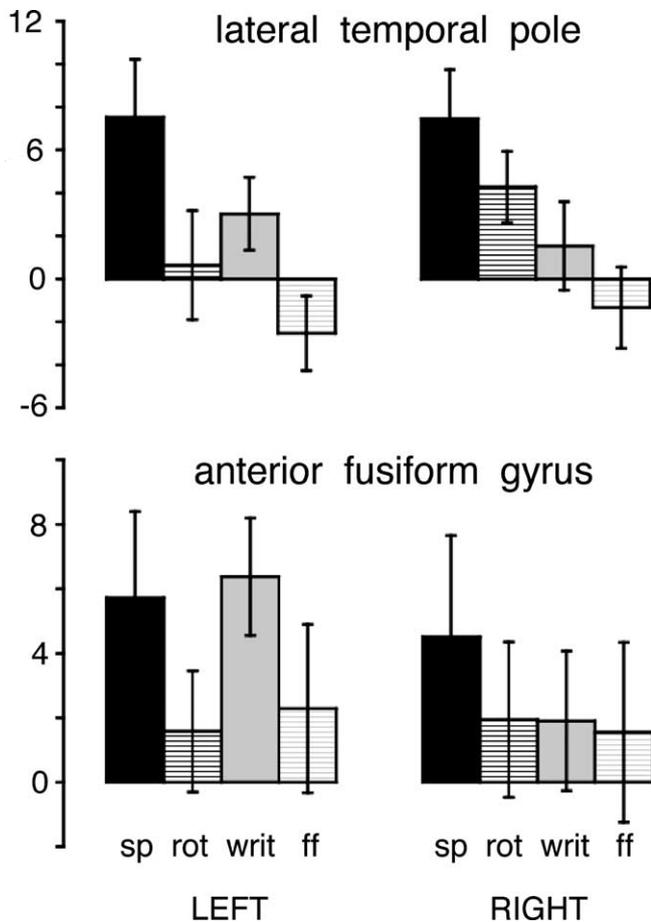


**Figure 5.** A demonstration on sagittal and coronal views, using the same presentation format as Figures 2 and 3, of activation common to the language conditions contrasted with their modality-specific unintelligible baseline conditions (color-coded in green), and common to the language conditions separately contrasted with the number task (color-coded in red). Although there was overlap of the activations in the left lateral TP (1), the use of the number task as the baseline condition revealed activity in the anterior FG (2) and the left TOP junction (3).

necessarily exclusive, response to verbal intelligibility in distributed heteromodal regions of the left temporal lobe.

## Discussion

This study demonstrated that the sensory pathways engaged in the processing of spoken and written language converge at four

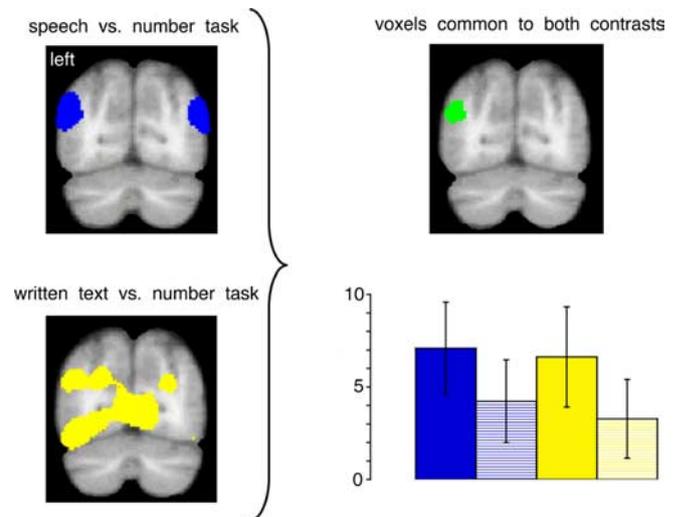


**Figure 6.** Histograms of mean estimated effect sizes (in arbitrary units), with 95% confidence intervals, in two anterior temporal ROIs in each hemisphere: the lateral TP and the anterior FG. The language conditions are shown as columns for speech (sp) and written text (writ). The unintelligible baseline conditions are shown as columns for rotated speech (rot) and false font (ff). The mean effect sizes for all four conditions are relative to the common baseline condition of the number task.

sites in the left cerebral hemisphere. These can be broadly classified into two groups according to the observed profile of responses across conditions: the STS and the lateral TP; and the anterior FG and TOP junction.

In the present study, much of the length of the STS was activated maximally by speech, with lesser activation observed for written text, which conveys phonetic information, and for rotated speech, which, although unintelligible, contains phonetic cues and features (Blessner, 1972). Previous functional imaging studies using auditory stimuli have related responses within or close to the left STS to prelexical phonetic processing (Scott et al., 2000; Liebenthal et al., 2005); and Ferstl and von Cramon (2001) demonstrated a response within anterior and posterior left STS in a contrast of written text with unpronounceable nonwords. Cross-modal integration of individual speech sounds and letters has also been demonstrated within the STS (van Atteveldt et al., 2004). The anatomical convergence of auditory and visual pathways enables the implicit (automatic) reciprocal processing of a word's sound and visual structure in literate subjects (Dijkstra et al., 1993). However, heteromodal responses within the left STS are not language-specific; the human STS also responds to non-verbal auditory and visual stimuli (Beauchamp et al., 2004).

Within the left STS, there was a stronger response to speech than rotated speech. Within the left lateral TP, there were signif-

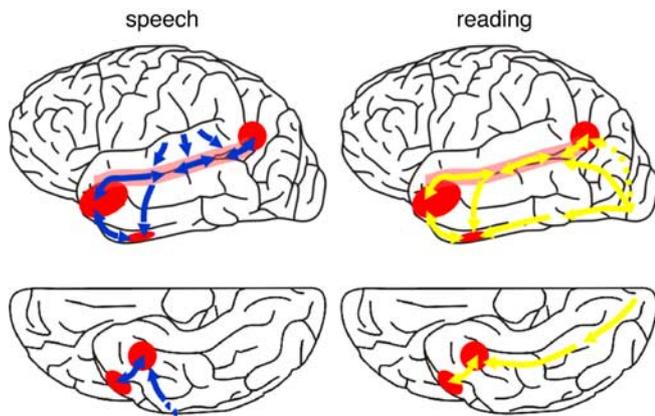


**Figure 7.** Activity at the TOP junction associated with speech (top left, in blue) and written text (bottom left, in yellow), each contrasted with the number task, coregistered onto a mean anatomical coronal magnetic resonance image, 72 mm posterior to the anterior commissure. There were voxels common to these two contrasts at the left TOP junction (top right, in green). Histograms of mean estimated effect sizes (expressed in arbitrary units), with 95% confidence intervals, in this region are shown (bottom right). The language conditions are shown as a blue column for speech and a yellow column for written text. Their unintelligible baseline conditions are shown as a blue-hatched column for rotated speech and a yellow-hatched column for false font. The mean effect sizes for all four conditions are relative to the common baseline condition of the number task.

icant responses to intelligible language in both modalities, regardless of the baseline condition used for comparison; however, responses were greater to auditory than visual stimuli. The lateral TP region incorporated several cytoarchitectonic regions, because no probabilistic atlas currently exists to separate them. Therefore, the combination of modality-independent responses to intelligible language and unimodal auditory responses observed in lateral TP is likely to reflect the inclusion of both ventral heteromodal areas and dorsal unimodal auditory association areas within this ROI.

Spoken and written language processing converged within the left TOP junction, where there was no effect of modality but where there was also activity in response to the unintelligible modality-specific baseline stimuli. The posterior region has been observed to respond to many behavioral conditions: to the “rest” or “passive” state, when self-generated thoughts may intrude (Binder et al., 1999; Gusnard and Raichle, 2001; Mazoyer et al., 2001); during the explicit manipulation of lexical semantic information (Binder et al., 2003); when establishing narrative coherence across sentences (Ferstl and von Cramon, 2002); and during the spontaneous generation of speech (Blank et al., 2002). The cumulative evidence indicates that this posterior region performs modality-independent and domain-general semantic processes across a wide range of internal and perceptual mental states.

Involvement of anterior basal temporal cortex in the processing of either spoken or written language has been demonstrated in previous functional imaging and electrophysiological studies (Nobre et al., 1994; Crinion et al., 2003; Scott et al., 2003; Sharp et al., 2004), which accords well with the modality-independent left anterior FG activation observed in the present study. The pattern of responses within the left anterior FG did not differ significantly from that within the TOP junction. However, because these two regions have very different patterns of afferent and efferent con-



**Figure 8.** An anatomical summary of the results, overlaid on the lateral (top row) and ventral (bottom row) surfaces of the left cerebral hemisphere, showing the convergence of spoken and written language in the left STS (pink shading) and in posterior (the TOP junction) and anterior (lateral TP and anterior FG) cortex (red shading). The auditory streams of processing are shown in blue, and the visual streams are in yellow. In the nonhuman primate, the heteromodal cortex of the left STS is strongly interconnected along its length (Padberg et al., 2003), and we have assumed that the same is true for the human STS. A direct connection between the anterior STS and inferotemporal cortex is suggested by anatomical connectivity data from nonhuman primates (Saleem et al., 2000). The left TOP junction may receive direct projections from occipital cortex (dashed yellow line) or indirectly, via the STS.

nections, it is unlikely that they perform functionally identical roles in the processing of language.

Convergent responses to intelligible language were located in the left hemisphere. Greater activity in the right temporal lobe in response to speech compared with reading can be attributed to the considerable nonverbal information carried by speech. Nevertheless, some clinical studies attest to a role for the right anterior temporal lobe in semantic processing (Lambon-Ralph et al., 2001), although neural reorganization in response to acute or progressive focal cortical damage may render the functional anatomy in patients quite different from that in normal subjects. This qualification also applies to the major argument advanced against roles for the lateral TP and the anterior FG in language comprehension; namely, the absence of a significant language deficit after anterior temporal lobectomy (ATL) for intractable temporal lobe epilepsy (TLE) (Hermann and Wyler, 1988; Davies et al., 1995). However, abnormal electrical activity, both ictal and interictal, is linked to reorganization of language cortex in TLE (Janszky et al., 2003). Furthermore, metabolic PET scans in TLE patients have demonstrated widespread interictal temporal lobe hypometabolism (Ryvlin et al., 1991), and altered anterior and inferior temporal lobe neurotransmitter function (Merlet et al., 2004). Therefore, the absence of significant comprehension deficits after ATL is likely to reflect a preoperative reorganization of the language system in the presence of left-lateralized TLE rather than an absence of language function in this region in the normal brain.

Figure 8 summarizes our results, by representing the processing streams from unimodal auditory and visual cortex to the left lateral TP, anterior FG, and TOP junction. This anatomical model contrasts sharply with those recently presented by Hickok and Poeppel (2004) and by Catani and ffytche (2005), which ascribe an exclusive role for posterior temporal or inferior parietal cortex in the access to verbal meaning. Hickok and Poeppel (2004) have speculated that the posterior region is responsible for lexical semantic processing and that activation of anterior temporal cortex by language stimuli might reflect syntactic process-

ing. However, this hypothesis finds little support from clinical studies of semantic dementia, in which anterior and inferior temporal lobe atrophy is associated with progressive impairment of verbal semantics but preserved syntactic knowledge (Hodges et al., 1992; Gorno-Tempini et al., 2004). In addition, functional imaging studies of grammatical processing have produced results with little anatomical consistency (Kaan and Swaab, 2002).

We have not included frontotemporal connections in the summary figure (Fig. 8). Although previous imaging studies investigating both spoken and written language processing (Chee et al., 1999; Booth et al., 2002) and cross-modal priming study (Buckner et al., 2000) have demonstrated inferior frontal gyrus activation, these studies involved performance of explicit metalinguistic tasks. In contrast, the present study and previous PET studies of implicit language processing (Scott et al., 2000; Crinion et al., 2003) do not emphasize a role for the left inferior frontal gyrus in implicit language comprehension. Therefore, although frontotemporal connections are undoubtedly important for a number of aspects of language processing, data from the present study of implicit language comprehension do not allow us to make inferences about these connections.

We conclude that widely separated anterior, inferior, and posterior regions in the left temporal lobe connect the unimodal processing streams of spoken and written language to distributed neocortical stores of attributional knowledge by which we “know” a particular object or concept (McClelland and Rogers, 2003). It is informative to relate the convergence of spoken and written language in the human brain to auditory and visual convergence in the monkey brain (Poremba et al., 2003). The dorsal bank of the monkey STS, termed the superior temporal polysensory area, is heteromodal cortex. The monkey inferior parietal lobe is also heteromodal, and the expansion of this region in the human, relative to other species, led Geschwind to propose that it plays a key role in human cognitive processing, including language (Geschwind, 1965); the area we term the TOP junction is located just ventral to the angular gyrus in the posterior part of the inferior parietal lobe. Of the other two areas responsive to language intelligibility, the lateral part of the anterior temporal lobe may represent an expansion of high-order heteromodal cortex from the anterior STS in the human compared with the monkey. It is more difficult to be certain about homology between the human and the monkey inferotemporal cortex. The monkey has no distinct fusiform gyrus (Gloor, 1997). Based on the anatomical evidence currently available (Insausti et al., 1998), language-related activity was located within the human homolog of the nonhuman primate cortical area TE, which is high-order visual association cortex. A proportion of area TE neurons in the monkey show context-dependent responses to auditory as well as visual stimuli (Gibson and Maunsell, 1997); and area TE and the more medial heteromodal paralimbic area (perirhinal cortex) encode learned associations between paired visual tokens (Naya et al., 2003). We conclude that, in the adult literate human brain, neurons in the anterior FG undergo organization, as the result of behavioral experience, to map between auditory and visual linguistic tokens (words) and stored semantic memories about their real-world referents.

In nonhuman primates, evidence suggests that, although anterior and posterior temporal lobe regions differ in their afferent and efferent cortical connections (for example, Petrides and Pandya, 2002) and therefore are unlikely to be functionally identical, they are nevertheless anatomically interconnected and probably functionally interdependent. Extrapolation from evidence in nonhuman primates (Padberg et al., 2003) suggests that there are

likely to be reciprocal connections along the length of the human STS; and infarction of the posterior temporal lobe has been shown to have functional effects on intact anterior temporal cortex (Crinion et al., 2006). Although, in the normal brain, functional separation between these regions may be driven by context-specific mental manipulation of language, a pathological loss of function in one region may, at least in part, be compensated for by the function of the others. The demonstration of multiple modality-independent sites within the normal language comprehension system opens a new perspective on possible processes of language recovery after brain injury.

## References

- Alain C, Arnott SR, Hevenor S, Graham S, Grady CL (2001) "What" and "where" in the human auditory system. *Proc Natl Acad Sci USA* 98:12301–12306.
- Alexander MP, Hiltbrunner B, Fischer RS (1989) Distributed anatomy of transcortical sensory aphasia. *Arch Neurol* 46:885–892.
- Aston JA, Gunn RN, Hinz R, Turkheimer FE (2005) Wavelet variance components in image space for spatiotemporal neuroimaging data. *NeuroImage* 25:159–168.
- Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004) Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat Neurosci* 7:1190–1192.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Rao SM, Cox RW (1999) Conceptual processing during the conscious resting state: a functional MRI study. *J Cogn Neurosci* 11:80–93.
- Binder JR, McKiernan KA, Parsons ME, Westbury CF, Possing ET, Kaufman JN, Buchanan L (2003) Neural correlates of lexical access during visual word recognition. *J Cogn Neurosci* 15:372–393.
- Blank SC, Scott SK, Murphy K, Warburton E, Wise RJ (2002) Speech production: Wernicke, Broca and beyond. *Brain* 125:1829–1838.
- Blessner B (1972) Speech perception under conditions of spectral transformation. I. Phonetic characteristics. *J Speech Hear Res* 15:5–41.
- Booth JR, Burman DD, Meyer JR, Gitelman DR, Parrish TB, Mesulam MM (2002) Modality independence of word comprehension. *Hum Brain Mapp* 16:251–261.
- Bozeat S, Lambon Ralph MA, Patterson K, Garrard P, Hodges JR (2000) Non-verbal semantic impairment in semantic dementia. *Neuropsychologia* 38:1207–1215.
- Brett M, Anton JL, Valabregue R, Poline JB (2002) Region of interest analysis using an SPM toolbox. *NeuroImage* 16:497.
- Buckner RL, Koutstaal W, Shacter DL, Rosen BR (2000) Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. *Brain* 123:620–640.
- Catani M, ffytche DH (2005) The rises and falls of disconnection syndromes. *Brain* 128:2224–2239.
- Chan D, Fox NC, Scallan RI, Crum WR, Whitwell JL, Leschziner G, Rossor AM, Stevens JM, Cipolotti L, Rossor MN (2001) Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. *Ann Neurol* 49:433–442.
- Chee MWL, O'Craven KM, Bergida R, Rosen BR, Savoy RE (1999) Auditory and visual word processing studied with fMRI. *Hum Brain Mapp* 7:15–28.
- Cohen L, Dehaene S (2004) Specialization within the ventral stream: the case for the visual word form area. *NeuroImage* 22:466–476.
- Crinion JT, Lambon Ralph MA, Warburton EA, Howard D, Wise RJ (2003) Temporal lobe regions engaged during normal speech comprehension. *Brain* 126:1193–1201.
- Crinion JT, Warburton EA, Lambon Ralph MA, Howard D, Wise RJ (2006) Listening to narrative speech after aphasic stroke: the role of the left anterior temporal lobe. *Cereb Cortex*, in press.
- Davies KG, Maxwell RE, Beniak TE, Destafney E, Fiol ME (1995) Language function after temporal lobectomy without stimulation mapping of cortical function. *Epilepsia* 36:130–136.
- Devlin JT, Russell RP, Davis MH, Price CJ, Moss HE, Fadili MJ, Tyler LK (2000) Susceptibility-induced loss of signal: comparing PET and fMRI on a semantic task. *NeuroImage* 11:589–600.
- Dijkstra T, Frauenfelder UH, Schreuder R (1993) Bidirectional grapheme-phoneme activation in a bimodal detection task. *J Exp Psychol Hum Percept Perform* 19:931–950.
- Ferstl EC, von Cramon DY (2001) The role of coherence and cohesion in text comprehension: an event-related fMRI study. *Cogn Brain Res* 11:325–340.
- Ferstl EC, von Cramon DY (2002) What does the frontomedian cortex contribute to language processing: coherence or theory of mind? *NeuroImage* 17:1599–1612.
- Friston KJ, Price CJ, Fletcher P, Moore C, Frackowiak RS, Dolan RJ (1996) The trouble with cognitive subtraction. *NeuroImage* 4:97–104.
- Geschwind N (1965) Disconnexion syndromes in animals and man. Part I. *Brain* 88:237–294.
- Gibson JR, Maunsell JH (1997) Sensory modality specificity of neural activity related to memory in visual cortex. *J Neurophysiol* 78:1263–1275.
- Gloor P (1997) The temporal lobe and limbic system. New York: Oxford UP.
- Gorno-Tempini ML, Dronkers NF, Rankin KP, Ogar JM, Phengrasamy L, Rosen HJ, Johnson JK, Weiner MW, Miller BL (2004) Cognition and anatomy in three variants of primary progressive aphasia. *Ann Neurol* 55:335–346.
- Gusnard D, Raichle ME (2001) Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci* 2:685–694.
- Hammers A, Allom R, Koeppe MJ, Free SL, Myers R, Lemieux L, Mitchell TN, Brooks DJ, Duncan JS (2003) Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp* 19:224–247.
- Hermann BP, Wyler AR (1988) Effects of anterior temporal lobectomy on language function: a controlled study. *Ann Neurol* 23:585–588.
- Hickok G, Poeppel D (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92:67–99.
- Hodges JR, Patterson K, Oxbury S, Funnell E (1992) Semantic dementia: progressive fluent aphasia with temporal lobe atrophy. *Brain* 115:1783–1806.
- Insausti R, Juottonen K, Soininen H, Insausti AM, Partanen K, Vainio P, Laakso MP, Pitkanen A (1998) MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *Am J Neuroradiol* 19:659–671.
- Janszky J, Jokeit H, Heinemann D, Schulz R, Woermann FG, Ebner A (2003) Epileptic activity influences the speech organization in medial temporal lobe epilepsy. *Brain* 126:2043–2051.
- Jernigan TL, Gamst AC, Fennema-Notestine C, Ostergaard AL (2003) More "mapping" in brain mapping: statistical comparison of effects. *Hum Brain Mapp* 19:90–95.
- Kaan E, Swaab TY (2002) The brain circuitry of syntactic comprehension. *Trends Cogn Sci* 6:350–356.
- Kertesz A, Sheppard A, MacKenzie R (1982) Localization in transcortical sensory aphasia. *Arch Neurol* 39:475–478.
- Lambon-Ralph MA, McClelland JL, Patterson K, Galton CJ, Hodges JR (2001) No right to speak? The relationship between object naming and semantic impairment: neuropsychological evidence and a computational model. *J Cogn Neurosci* 13:341–356.
- Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA (2005) Neural substrates of phonemic perception. *Cereb Cortex* 15:1621–1631.
- Marinkovic K, Dhond RP, Dale AM, Glessner M, Carr V, Halgren E (2003) Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron* 38:487–497.
- Mazoyer B, Zago L, Mellet E, Bricogne S, Etard O, Houde O, Crivello F, Joliot M, Petit L, Tzourio-Mazoyer N (2001) Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Res Bull* 54:287–298.
- McClelland JL, Rogers TT (2003) The parallel distributed processing approach to semantic cognition. *Nat Rev Neurosci* 4:310–322.
- Merlet I, Ryvlin P, Costes N, Dufournel D, Isnard J, Faillenot I, Ostrowsky K, Lavenne F, Le Bars D, Maugeyrie F (2004) Statistical parametric mapping of 5-HT<sub>1A</sub> receptor binding in temporal lobe epilepsy with hippocampal ictal onset on intracranial EEG. *NeuroImage* 22:886–896.
- Mesulam MM (1998) From sensation to cognition. *Brain* 121:1013–1052.
- Naya Y, Yoshida M, Miyashita Y (2003) Forward processing of long-term associative memory in monkey inferotemporal cortex. *J Neurosci* 23:2861–2871.
- Nichols T, Brett M, Andersson J, Wagner T, Poline J-B (2005) Valid conjunction inference with the minimum statistic. *NeuroImage* 25:653–660.
- Nobre AC, Allison T, McCarthy G (1994) Word recognition in the human inferior temporal lobe. *Nature* 372:260–263.

- Padberg J, Seltzer B, Cusick CG (2003) Architectonics and cortical connections of the upper bank of the superior temporal sulcus in the rhesus monkey: an analysis in the tangential plane. *J Comp Neurol* 467:418–434.
- Petrides M, Pandya DN (2002) Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and cortico-cortical connection patterns in the monkey. *Eur J Neurosci* 16:291–310.
- Poremba A, Saunders RC, Crane AM, Cook M, Sokoloff L, Mishkin M (2003) Functional mapping of the primate auditory system. *Science* 299:568–572.
- Rauschecker JP, Tian B (2000) Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proc Natl Acad Sci USA* 97:11800–11806.
- Ryvlin P, Cinotti L, Froment JC, Le Bars D, Landais P, Chaze M, Galy G, Lavenne F, Serra JP, Mauguiere F (1991) Metabolic patterns associated with non-specific magnetic resonance imaging abnormalities in temporal lobe epilepsy. *Brain* 114:2363–2383.
- Saleem KS, Suzuki W, Tanaka K, Hashikawa T (2000) Connections between anterior inferotemporal cortex and superior temporal sulcus regions in the macaque monkey. *J Neurosci* 20:5083–5101.
- Scott SK, Johnsrude IS (2003) The neuroanatomical and functional organization of speech perception. *Trends Neurosci* 26:100–107.
- Scott SK, Blank SC, Rosen S, Wise RJ (2000) Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123:2400–2406.
- Scott SK, Leff AP, Wise RJ (2003) Going beyond the information given: a neural system supporting semantic interpretation. *NeuroImage* 19:870–876.
- Sharp DJ, Scott SK, Wise RJ (2004) Monitoring and the controlled processing of meaning: distinct prefrontal systems. *Cereb Cortex* 14:1–10.
- Stark CEL, Squire LR (2001) When zero is not zero: the problem of ambiguous baseline conditions in fMRI. *Proc Natl Acad Sci USA* 98:12760–12766.
- Thioux M, Pesenti M, Costes N, De Volder A, Seron X (2005) Task-independent semantic activation for numbers and animals. *Cogn Brain Res* 24:284–290.
- Turkheimer FE, Brett M, Aston JAD, Leff AP, Sargent PA, Wise RJS, Grasby PM, Cunningham VJ (2000) Statistical modeling of positron emission tomography images in wavelet space. *J Cereb Blood Flow Metab* 20:1610–1618.
- Ungerleider LG, Haxby JV (1994) “What” and “where” in the human brain. *Curr Opin Neurobiol* 4:157–165.
- Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: *Analysis of visual behavior* (Dingle DJ, Goodale MA, Mansfield RJW, eds), pp 549–586. Cambridge, MA: MIT.
- van Atteveldt N, Formisano E, Goebel R, Blomert L (2004) Integration of letters and speech sounds in the human brain. *Neuron* 43:271–282.