Dissociating Verbal and Nonverbal Conceptual Processing in the Human Brain

Guillaume Thierry[1] and Cathy J. Price[2]

Abstract

■ Functional neuroimaging has highlighted a left-hemisphere conceptual system shared by verbal and nonverbal processing despite neuropsychological evidence that the ability to recognize verbal and nonverbal stimuli can doubly dissociate in patients with left- and right-hemisphere lesions, respectively. Previous attempts to control for perceptual differences between verbal and nonverbal stimuli in functional neuroimaging studies may have hidden differences arising at the conceptual level. Here we used a different approach and controlled for perceptual confounds by looking for amodal verbal and nonverbal conceptual activations that are common to both the visual and auditory modalities. In addition to the left-hemisphere conceptual system activated by all meaningful stimuli, we observed the left/right double dissociation in verbal and nonverbal conceptual processing, predicted by neuropsychological studies. Left middle and superior temporal regions were selectively more involved in comprehending words—heard or read—and the right midfusiform and right posterior middle temporal cortex were selectively more involved in making sense of environmental sounds and images. Thus, the neuroanatomical basis of a verbal/nonverbal conceptual processing dissociation is established. ■

INTRODUCTION

Neuroimaging investigations targeting the neural basis of conceptual processing have highlighted the involvement of a shared left-lateralized network in both the verbal and nonverbal domains. By conceptual processing, we mean the access to long-term memory representations of meaningful objects (concrete or abstract), the retrieval of their functional properties, and the mechanisms by which these properties are assessed/compared/related prior to a decision. In the verbal domain, left-lateralized activation has been identified for tasks such as categorization (Binder et al., 1997; Démonet, Chollet, et al., 1992), comprehension (Scott, Blank, Rosen, & Wise, 2000), rehearsal (Paulesu, Frith, & Frackowiak, 1993), selection (Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997), and reading for meaning (for reviews, see Démonet, Thierry, & Cardebat, 2005; Price, 2000). Likewise, in the nonverbal domain, left-lateralized activation has been identified for conceptual tasks with pictures of objects (Bright, Moss, & Tyler, 2004; Whatmough, Chertkow, Murtha, & Hanratty, 2002; Mummery, Ashburner, Scott, & Wise, 1999; Gorno-Tempini et al., 1998; Mummery, Patterson, Hodges, & Price, 1998; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996), environmental sounds (Thierry, Giraud, & Price, 2003; Giraud & Price, 2001; Humphries, Willard, Buchsbaum, & Hickok, 2001; Engelien et al., 1995), and during the monitoring of harmonic structures (Levitin &

Menon, 2003; Maess, Koelsch, Gunter, & Friederici, 2001). The existence of a shared left-hemispheric system involved in verbal and nonverbal conceptual processing is also supported by voxel-based reconstructions in aphasic patients who sustain deficits in both speech comprehension and environmental sound recognition (see also Saygin, Dick, Wilson, Dronkers, & Bates, 2003; Clarke, Bellmann, Meuli, Assal, & Steck, 2000; Schnider, Benson, Alexander, & Schnider-Klaus, 1994).

In addition to the shared left-lateralized conceptual system, there is also evidence for verbal/nonverbal functional dissociations at both the perceptual and conceptual levels. Patients with pure word deafness (Seliger et al., 1991) and deep dyslexics (Coltheart, 1980), for instance, are impaired on auditory or visual verbal tasks but can perform normally on nonverbal tasks. Reciprocally, there are reports of patients with auditory agnosia without aphasia (Fujii et al., 1990; Spreen, Benton, & Fincham, 1965) or visuoperceptual deficits (Warrington, 1982), who show nonverbal impairments but achieve normal levels of verbal comprehension on spoken and written language tasks, respectively.

Significant performance differences in accessing the meaning of pictures as compared to words (e.g., Shallice, 1993; McCarthy & Warrington, 1988; Shallice, 1988; Warrington & Shallice, 1984) have led to the hypothesis that there are multiple conceptual stores and that verbal and nonverbal access to meaning may be neuroanatomically segregated (e.g., Warrington & Crutch, 2004; Coltheart et al., 1998; McCarthy & Warrington,

[1]University of Wales Bangor, UK, [2]University College London, UK

1988; Beauvois & Saillant, 1985). The evidence is primarily drawn from the visual modality because ad hoc neurological syndromes (e.g., transcortical sensory aphasia [Lichtheim, 1885] and semantic refractory access dysphasia [see Warrington & Crutch, 2004]) are characterized by severe auditory comprehension deficits incompatible with auditory verbal testing. However, even in the visual modality, the level (perceptual or conceptual) at which the dissociations occur is often difficult to determine on the basis of neuropsychological testing because disruption at the perceptual level will have conceptual repercussions. Therefore, the multiplicity of conceptual stores remains highly debated (e.g., Lambon Ralph, Graham, Patterson, & Hodges, 1999; Caramazza, Hillis, Rapp, & Romani, 1990; Riddoch, Humphreys, Coltheart, & Funnel, 1988; for a review, see Saffran & Schwartz, 1994).

Functional imaging evidence for a verbal/nonverbal dissociation is implied when the results from verbal and nonverbal activation studies are compared in either the auditory modality (von Kriegstein & Giraud, 2004; von Kriegstein, Eger, Kleinschmidt, & Giraud, 2003; Zatorre, Belin, & Penhune, 2002; Zatorre & Belin, 2001; Belin, Zatorre, Lafaille, Ahad, & Pike, 2000;) or the visual modality (Cohen, Lehericy, et al., 2002; Trojano et al., 2002; Cohen, Dehaene, et al., 2000; Kanwisher, McDermott, & Chun, 1997; Smith, Jonides, & Koeppe, 1996; Sergent, Ohta, & MacDonald, 1992). However, no clear double dissociation between verbal and nonverbal conceptual processing has yet been demonstrated when perceptual differences are controlled (e.g., by subtracting activations elicited by meaningless, scrambled stimuli). For example, the study by Vandenberghe, Price, et al. (1996) controlled for perceptual confounds between words and pictures by comparing conceptual tasks (e.g., Which of these objects is the biggest in real life?) to perceptual tasks (e.g., Which of these objects is the biggest on the screen?) on one type of stimulus relative to the other (i.e., by characterizing a stimulus by task interaction). The difficulty with this approach is that it excludes verbal or nonverbal conceptual processing that may have occurred irrespective of task. Thus, to fully characterize differences between verbal and nonverbal processing, we propose that three conditions are met: (i) Verbal and nonverbal stimuli must be compared directly to identify differences at all levels of processing; (ii) tasks performed on verbal and nonverbal stimuli must be matched in structure, cognitive requirements, and difficulty; (iii) modality- and task-independent verbal and nonverbal differences must be established.

In a previous neuroimaging study involving normal volunteers (Thierry et al., 2003), we directly compared processing of spoken words (e.g., ''cow mooing'') and environmental sounds (e.g., cow mooing) matched for meaning using semantic categorization and sequence interpretation tasks. We controlled for some of the low-level perceptual differences between verbal and nonverbal stimuli by using baseline conditions that involved unintelligible noise bursts created by scrambling each speech and environmental sound stimulus. A shared, left-lateralized conceptual system was observed for both types of stimuli but, in addition, left anterior superior temporal activation was greater for spoken words than sounds and, conversely, right posterior superior temporal activation was greatest for sounds. More importantly, this functional dissociation was not observed for words and sounds during repetition and naming (Giraud & Price, 2001), as indicated by a significant interaction between task and stimulus type when data from both studies were combined. Although conceptual processing is likely to take place in repetition and naming tasks, the nature of conceptual operations is uncontrolled, whereas in the Thierry et al. (2003) study, conceptual requirement of the tasks was equated for spoken words and environmental sounds. We therefore concluded that differences in the Thierry et al. study did not reflect perceptual differences between verbal and nonverbal sources and instead proposed that they arose at the level of accessing meaning (i.e., at a conceptual level).

In the current study, we conducted a visual experiment to establish verbal/nonverbal dissociations that occur irrespective of sensory modality. By combining data from this new experiment with that from the auditory version previously reported (Thierry et al., 2003), we were able to determine whether the verbal versus nonverbal dissociation was seen in both stimulus modalities, or in the auditory modality only. In other words, the aim was to dissociate amodal from modality-specific verbal/nonverbal dissociations.

In the auditory version of the experiment, 12 participants listened to sequences of environmental sounds and spoken words. In the visual version of the experiment, another 12 participants viewed mute videos and sequences of written words (Figure 1). All four sets of stimuli were matched for meaning and interpretability. On the basis of the neuropsychological evidence we predicted that (a) verbal processing (common to spoken words and text displays) as compared to nonverbal processing (common to environmental sounds and mute videos) would yield greater activation in left superior temporal regions (Thierry et al., 2003; Scott et al., 2000) and (b) nonverbal as compared to verbal processing would yield greater activation in the right hemisphere (Thierry et al., 2003; Coltheart, 1980).

METHODS

Participants

Twenty-four native speakers of English (mean age = 26.3 ± 8.4 years, all men) gave written consent to participate in 12 positron emission tomography (PET) scans (Siemens CTI III camera) involving intravenous
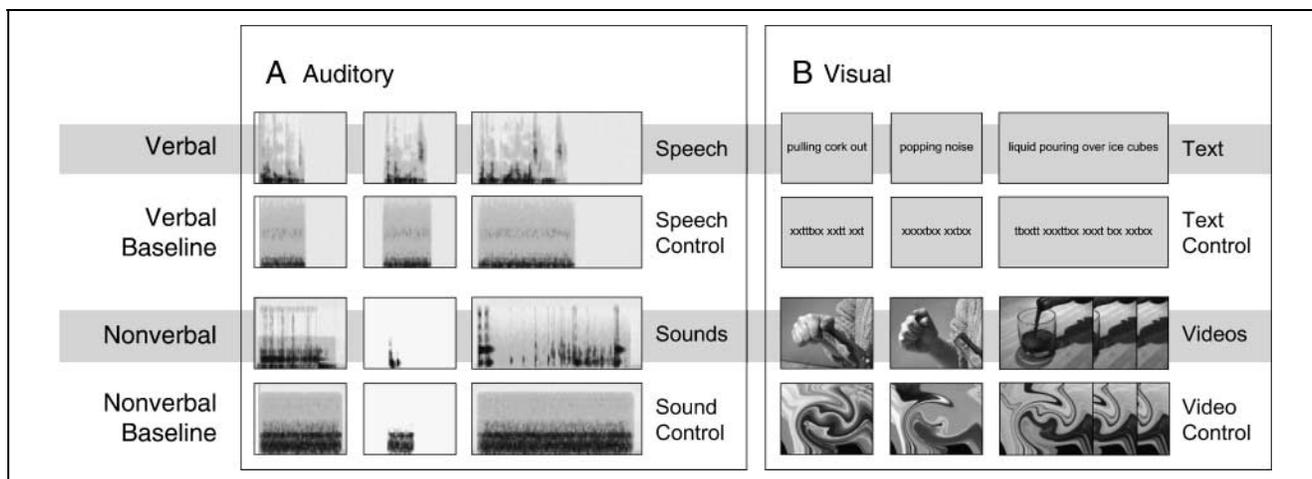
Figure 1. Verbal and nonverbal stimulus types used in the auditory and visual components of the experiment. (A) In the auditory version, semantic decisions on speech samples (speech) were contrasted to equivalent decisions on environmental sound samples (sounds). The auditory baseline entailed meaningless sequences generated for each meaningful sequence by scrambling words (speech control) or sounds (sound control). (B) In the visual version, sequences of written words (text) were contrasted to mute video clips (videos). The visual baseline entailed pseudorandom strings of Xs and Ts (text control) and the mute videos passed through distortion filters making the videos meaningless (video control).

injection of water labeled with ^{15}O. The dose received was 9 mCi per measurement. The study was approved by the joint ethics committee of the Institute of Neurology (University College London [UCL]) and the National Hospital for Neurology and Neurosurgery (UCLH NHS Trust) and the UK Administration of Radioactive Substances Advisory Committee (ARSAC). PET experiments on healthy individuals cannot include women of child-bearing age. All subjects were strongly right-handed, wrote with their right hand, kicked with their right foot, and had no family history of left-handedness.

Design

Twelve participants were presented with the auditory stimuli (Experiment 1) and the other 12 were presented with visual stimuli (Experiment 2). In both experiments, participants were exposed to verbal stimuli (speech or text) and nonverbal stimuli (environmental sounds or mute video clips) and had to perform two different conceptual tasks (categorization or sequence interpretation) on each stimulus type; see below for details. Activation conditions therefore conformed to a fully balanced 2 (between subjects) × 2 × 2 (within subject) design with two stimulus modalities (auditory and visual), two types of stimuli (verbal and nonverbal), and two tasks (categorization and sequence interpretation).

Stimuli

The activation stimuli were verbal or nonverbal sequences lasting 17 sec, on average (range 15–20 sec),

and ending with a distinctive signal (200-msec beep in the auditory conditions and white disk flashed for 200 msec at the center of the display in the visual conditions). Auditory sequences were made of digitized spoken words (e.g., drink story: ''pulling cork out ... popping noise ... pouring liquid over ice cubes ... someone sipping ... someone sighing'') and digitized environmental sounds. The phrases used provided minimal syntactic information (determinants were systematically excluded) and excluded superfluous semantic information (e.g., adjectives) to minimize differences between speech and sound conditions. Environmental sounds are considered a good nonverbal counterpart of human speech because they are auditory, they have a complex spectral structure, their duration can be adjusted to match word duration, and common sounds are easily identifiable (see Thierry et al., 2003). Visual sequences consisted of written words and mute video clips. Written word sequences were the direct transcription of the auditory phrases. Video clips were preferred to static images to avoid ambiguity in interpreting actions and to mimic the temporal structure of the spoken word/environmental sound sequences used in the auditory tasks. Sequences were matched within and across modality as closely as possible in terms of meaning, number of events, rhythm, and duration (Figure 1). Fifty percent of sequences included a stimulus that referred to an animal and 50% were logically ordered (events occurring in the most expected order). An unintelligible control sequence was generated for each meaningful one: Sound files (individual spoken words and sounds) were scrambled by using a random splicing procedure (Thierry et al., 2003); letters of the written words were replaced by pseudorandom strings of Xs

and *T*s; and videos were distorted using combined polar coordinates and twirling transformations (Figure 1). None of these baseline conditions contained any recognizable stimulus. Baselines were included to control for low-level perceptual differences between verbal and nonverbal contexts. Eight out of 14 original sequences yielding comparable behavioral performance were selected on the basis of a pilot behavioral screening involving 15 participants in the auditory version and 12 in the visual version (data not reported).

Tasks

Participants performed two tasks while either listening to words and sounds or viewing text and images. In a categorization task, they indicated with a key press whether a reference to an animal was present within each sequence (animal/no animal). This task was selected because conceptual categorization is equally achievable in a verbal and nonverbal context and it is easy. In a sequence interpretation task, participants indicated whether each sequence was logically ordered (ordered/disordered). For instance, when the event ''someone sighing'' was presented before ''someone sipping'' at the end of the drink story, the trial required a ''disordered'' judgment. This task was selected because it is equally achievable in a verbal and nonverbal context but is ambiguous and difficult. It was therefore more challenging and placed additional demands on attention and working memory. Any interaction between condition (verbal vs. nonverbal) and task (categorization vs. sequence interpretation) would indicate a modulation by difficulty and/or verbalization strategies. In the baseline conditions (scrambled speech, scrambled sounds, pseudotext displays, or distorted videos), participants pressed a key at the end of the meaningless sequence to control for finger movement.

Procedure

PET scans involved eight activation scans (verbal categorization, nonverbal categorization, verbal sequence interpretation, and nonverbal sequence interpretation) and four baseline scans (meaningless sequences derived by scrambling each verbal and nonverbal stimulus). Each participant heard/viewed a total of 16 different sequences (4 per scan) for both the verbal and the nonverbal conditions. The order of scans was counterbalanced over participants; the order of sequences within scans was pseudorandomized; and the order of stimuli within each sequence was never repeated. In all conditions, participants were instructed to respond as quickly and accurately as possible with a mouse button press after seeing the signal ending a sequence. Finger responses were alternated within subjects across blocks, and behavioral data were recorded online.

Data Processing

Data were acquired using a Siemens/CPS EXACT HR+ (model 962) PET scanner (Siemens/CTI, Knoxville, TN). Realignment of images, normalization, and statistics were performed with SPM99 (www.fil.ion.ucl.ac.uk/spm; Friston et al., 1995). Images were spatially smoothed with a 6-mm Gaussian filter. The statistical model partitioned the two groups of 12 subjects with six conditions per group. Summing over conceptual task (categorization and sequence interpretation), we computed the following effects:

(A) Auditory modality:

> (a_1) spoken words > speech control;
> (a_2) environmental sounds > sound control;
> (a_3) spoken words > environmental sounds;
> (a_4) environmental sounds > spoken words; and
> (a_5) the interaction of (a_3) and (a_4) with conceptual task.

(B) Visual modality:

> (b_1) text displays > text control;
> (b_2) videos > video control;
> (b_3) text > videos;
> (b_4) videos > text; and
> (b_5) the interaction of (b_3) and (b_4) with conceptual task.

All contrasts were computed within subject because the focus was on effects that were common to auditory and visual modality, irrespective of sensory modality. The common semantic system was identified by the main effect of (a_1) + (a_2) + (b_1) + (b_2) inclusively masked with (a_1), (a_2), (b_1) and (b_2). Effects of verbal > nonverbal that were common to stimulus modality were identified by the main effect of (a_3) + (b_3) inclusively masked with (a_3), (b_3), (a_1) and (b_1). Effects of nonverbal > verbal that were common to stimulus modality were identified by the main effect of (a_4) + (b_4) inclusively masked with (a_4), (b_4), (a_2), and (b_2). For each effect, we also report the details of contributing contrasts (see Tables 1 and 2). The statistical threshold for significance was $p < .05$ after correction for multiple comparisons across the whole brain. The threshold for the inclusively masking was $p < .05$ uncorrected.

RESULTS

Behavioral Performance

Behavioral results are shown in Figure 2. In the auditory experiment, the main effect of task on reaction times was significant, $F(2,10) = 45.8, p < .0001$ (sequence interpretation > categorization > baseline), but there was

Table 1. Activation Common to All Conditions Relative to Stimulus-specific Baselines

Region	All Modalities	Auditory Words	Auditory Sounds	Visual Words	Visual Videos
L post STG	−56 −40 4 (6.1)	−56 −42 2 (5.6)	−58 −38 6 (3.4)	−60 −38 8 (2.9)	−54 −40 4 (3.4)
L post STS	−60 −38 −4 (4.8)	−60 −36 0 (5.6)	−58 −38 −10 (3.6)	−60 −38 −4 (3.0)	−56 −38 −8 (3.8)
L ant STS	−58 8 −16 (5.1)	−56 10 −16 (4.9)	−56 6 −16 (2.6)	−58 6 −16 (3.0)	−58 10 −20 (3.3)
L inf FG	−52 14 24 (5.0)	−54 10 22 (2.3)	−50 14 24 (4.4)	−50 14 22 (3.8)	−58 14 24 (2.7)
R cereb	+22 −86 −34 (4.8)	+22 −86 −28 (4.3)	+16 −88 −36 (3.2)	+24 −76 −32 (3.2)	+22 −86 −34 (3.2)

L = left; R = right; post = posterior; ant = anterior; inf = inferior; STG = superior temporal gyrus; STS = superior temporal sulcus; FG = frontal gyrus; cereb = cerebellum. Talairach coordinates are given in the order x, y, z. Z scores are given in parentheses.

no effect of stimulus type, $F(1,11) < 1$, and no interaction, $F(2,10) < 1$. Similarly, the main effect of task on error rates was significant, $F(2,10) = 37.6$, $p < .001$ (sequence interpretation > categorization > baseline), but there was no effect of stimulus type, $F(1,11) < 1$, and no interaction, $F(2,10) < 1$. In the visual modality, the pattern of differences was the same. The main effect of task on reaction times was significant, $F(2,10) = 17.6$,

$p < .001$ (sequence interpretation > categorization > baseline), but there was no effect of stimulus type, $F(1,11) < 1$, and no interaction, $F(2,10) < 1$. Similarly, the main effect of task on error rates was significant, $F(2,10) = 78.1$, $p < .001$ (sequence interpretation > categorization > baseline), but there was no effect of stimulus type, $F(1,11) = 2.5$, ns, and no interaction, $F(2,10) = 1.5$, ns.

Table 2. Verbal versus Nonverbal Comparisons in Visual and Auditory Modalities

Region	Over Modalities	Auditory	Visual
(a) Verbal > nonverbal			
L post sup TS	−62 −36 4 (5.3)	−60 −38 0 (4.7)	−62 −36 4 (4.0)
	−62 −28 −2 (4.9)	−56 −24 −6 (4.8)	−64 −30 −2 (4.8)
	−56 −6 2 (5.4)	−58 −12 2 (4.7)	−56 −6 2 (3.6)
	−62 −4 −8 (5.3)	−60 −8 −10 (5.1)	−62 −4 −6 (3.3)
L ant sup TS	−44 14 −22 (4.7)	−44 14 −22 (4.5)	−46 14 −22 (2.6)
L vent IFG	−46 24 −10 (4.7)	−46 26 −12 (2.4)	−48 22 −10 (5.1)
L vent MFG	$p > .05$	$p > .05$	−30 54 −4 (5.0)
(b) Nonverbal > verbal			
R. Fus	+46 −46 −22 (5.2)	+44 −48 −26 (3.5)	+46 −44 −20 (4.8)
R. sup TG	+60 −54 4 (5.0)	+60 −56 8 (3.7)	+60 −54 4 (4.7)
	+70 −34 14 (4.1)[#]	+70 −34 14 (3.6)	+72 −32 14 (2.5)
	+64 −20 8 (3.9)*	+60 −16 8 (5.0)	+60 −22 8 (3.3)
R inf TG/mid TG	+54 −70 2 (3.6)*	+54 −70 2 (2.2)	+54 −70 2 (6.5)
	ns	ns	+46 −68 −2 (6.1)
R mid T/mid Oj	ns	ns	+56 −68 10 (6.5)
L mid OG	ns	ns	−44 −82 8 (9.0)
L inf O	ns	ns	−42 −80 −8 (7.3)

Talairach coordinates are given in the order x, y, z. Z scores are given in parentheses. L = left; R = right; ant = anterior; mid = middle; post = posterior; vent = ventral; sup = superior; inf = inferior; T= temporal; F = frontal; O = occipital; Fus = fusiform; S = sulcus; G = gyrus; j = junction; ns = not significant ($p > .05$ uncorrected).

*$p < .05$ corrected for multiple comparisons ($Z > 4.7$); significant in one modality but not over stimulus modalities.

[#]Not significant after correction for multiple comparisons but consistent with Thierry et al. (2003).

Figure 2. Behavioral results. (A) Auditory experiment. (B) Visual experiment. Bars depict reaction times (left scale) and circles depict error rates (right scale). White bars depict performance for verbal material (speech or text) and gray bars correspond to nonverbal performance (sounds or videos). Error bars indicate standard errors in all cases.
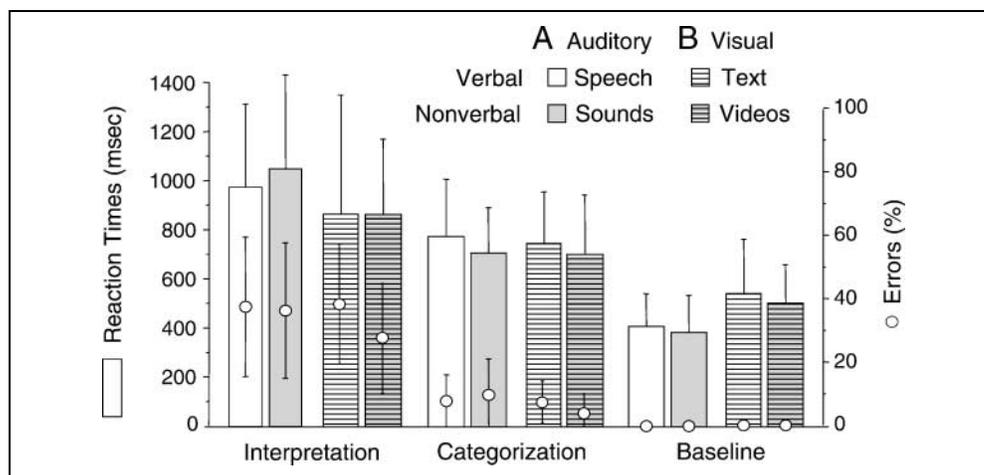
In sum, participants were slowest and made the largest number of errors in the sequence interpretation task; their performance in the categorization task was intermediate; and performance was optimal for the baseline condition independent of the sensory modality in which the stimuli were presented. Critically, there was no effect of stimulus type (verbal/nonverbal) overall and no interaction between stimulus type and task (i.e., no verbal/nonverbal difference within each of the tasks).

Neuroimaging Results

First, we identified activation common to verbal and nonverbal conceptual processing irrespective of sensory modality (Figure 3). We then focused on differential activation for verbal and nonverbal conditions (Figures 4 and 5). In all the results reported below, there was no significant interaction of condition (verbal vs. nonverbal or visual vs. auditory) with task (categorization vs. sequence interpretation). Therefore, all effects are reported after averaging over tasks.

Common Conceptual Activations

Consistent with previous studies (Vandenberghe, Price, et al., 1996), regions commonly activated by meaningful versus meaningless conditions independent of sensory modality and independent of verbal status revealed a shared network in the left hemisphere including the posterior superior temporal gyrus, the superior temporal sulcus (anterior and posterior), the left inferior

Figure 3. Brain activations common to all meaningful conditions. Activations for all meaningful conditions relative to baseline are shown in white (at $p < .05$ corrected for multiple comparisons) on (A) a surface model of the left and right hemispheres (L = left, R = right) and (B) sagittal and coronal brain slices. See Tables 1 and 2 for coordinates of activation peaks.
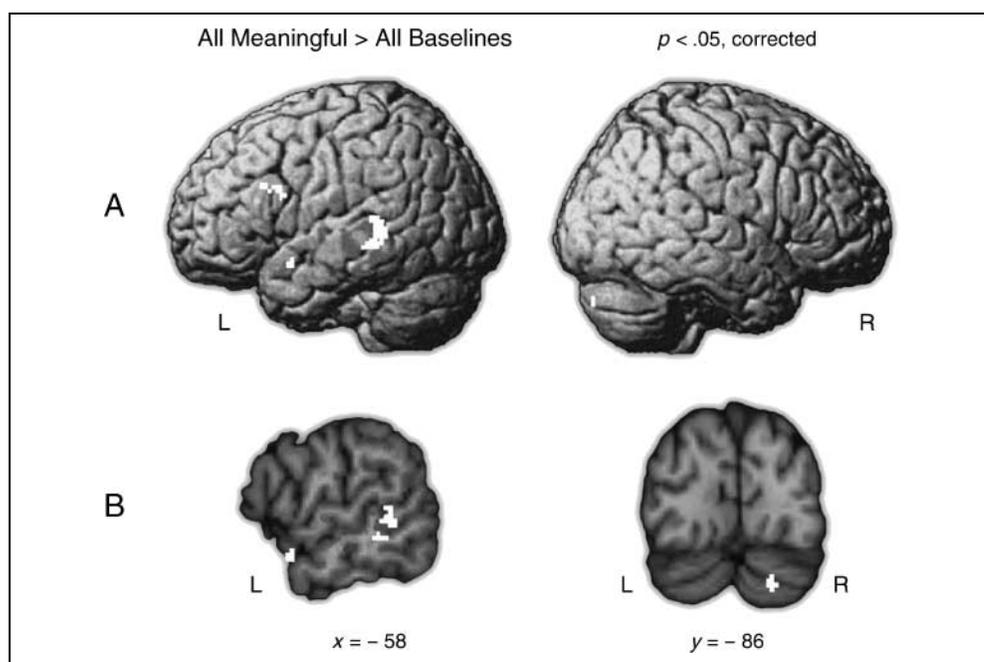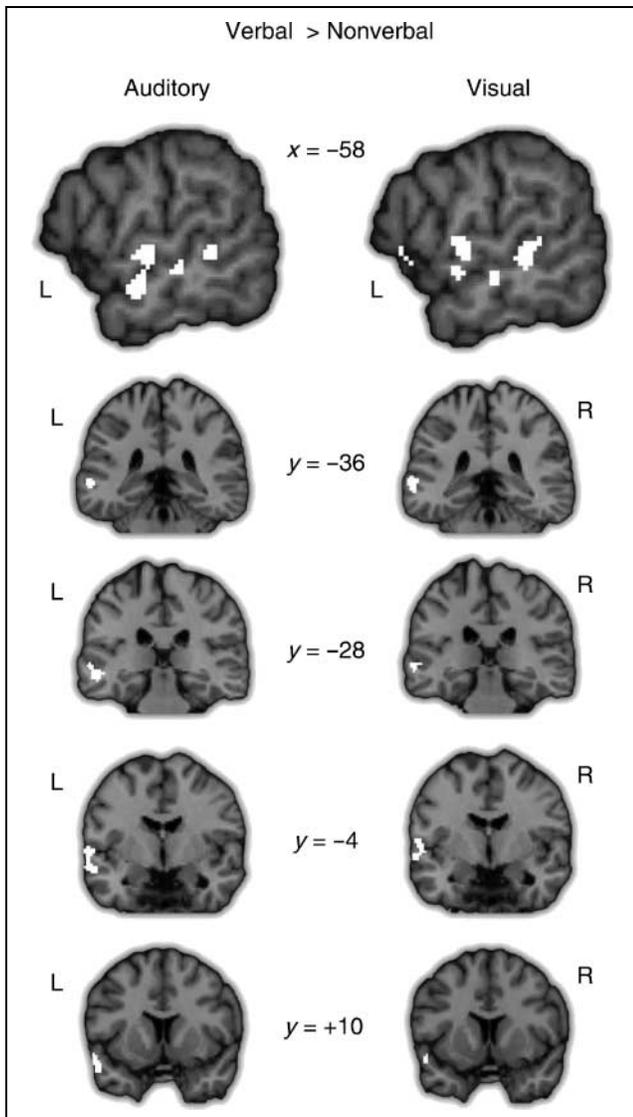
Figure 4. Greater activation for verbal than nonverbal conditions. All activations are shown in white. Top row shows the similarity between the activation for verbal versus nonverbal conditions in the auditory (left) and visual (right) modalities rendered on a sagittal slice ($x = -58$) of the MNI canonical brain. Second to fifth rows show the location of the different regions on coronal slices. L = left; R= right. The threshold is set at $p < .001$ uncorrected for the auditory modality and $p < .01$ uncorrected for the visual modality. Over modalities all effects were significant at $p < .05$ corrected for multiple comparisons. See Table 2 for details.

frontal gyrus, and the right cerebellum (see Table 1 and Figure 3).

Amodal Verbal/Nonverbal Dissociation

As predicted, differential activation for verbal relative to nonverbal conditions was observed in the left hemisphere and, reciprocally, greater activation for nonverbal relative to verbal conditions was observed in the right hemisphere. More importantly, these differences were observed within subjects irrespective of sensory modal-

ity or subject group. For verbal processing, activation was significantly greater than nonverbal processing in five regions along the left superior temporal sulcus and in the left ventral inferior frontal gyrus (Table 2a and Figure 4). For nonverbal processing, activation was significantly greater than verbal processing in the right midfusiform and the right posterior middle temporal gyrus (Table 2b and Figure 5). No other areas reached our criterion of significance ($p < .05$ corrected for the entire brain).

Modality-specific Verbal/Nonverbal Dissociations

There were also verbal versus nonverbal differences that were significant only in the visual modality (Table 2). A

Figure 5. Greater activation for nonverbal than verbal conditions. All activations are shown in white. Top row shows the similarity between the activation for verbal versus nonverbal conditions in the auditory (left) and visual (right) modalities rendered on a sagittal slice ($x = -58$) of the MNI canonical brain. Second to fourth rows show the location of the different regions on coronal slices. L= left; R = right. The threshold is set at $p < .001$ uncorrected for the auditory modality and $p < .01$ uncorrected for the visual modality. See Table 2 for details.

left ventral middle frontal region was activated by semantic decisions on visual words more than all other conditions, but this region was also weakly activated by all semantic conditions. This activation was therefore not specific to visual words. Several bilateral occipital and right middle/inferior temporal regions were selectively associated with semantic decisions on videos. However, as these effects were not observed in the other stimulus modalities, we cannot exclude the contribution of perceptual confounds within each modality.

Consistency with Previous Findings

Greater activation for verbal relative to nonverbal processing replicates that shown in the auditory modality by Thierry et al. (2003) and the visual modality by Vandenberghe, Price, et al. (1996). However, the activation common to the two sensory modalities revealed by the reverse contrast, that is, greater for nonverbal than verbal processing, has not been reported by previous studies including those by Thierry et al. and Vandenberghe, Price, et al. The amodal nonverbal activation described here is therefore a fundamentally new finding.

Thierry et al. (2003) did not report the nonverbal versus verbal activation in the right posterior middle temporal and right midfusiform regions because, in the auditory condition alone, these effects were below their critical threshold ($p < .05$ corrected for the entire brain); see Table 2b. However, when combined with the visual data, the conjoint effect of auditory and visual nonverbal processing becomes highly significant.

Vandenberghe, Price, et al. (1996) did not report the nonverbal versus verbal effect in the right posterior middle temporal and right midfusiform regions because they did not directly compare semantic decisions on pictures and words. Instead, they focused on the interaction of task (semantic vs. perceptual) and stimulus (pictures of objects vs. written words) in an attempt to tease apart semantic from perceptual differences. This approach will have excluded semantic activation that occurred irrespective of task. We therefore reanalyzed the Vandenberghe et al. data and directly contrasted semantic decisions on pictures of objects and semantic decisions on written names. Consistent with the current results, the Vandenberghe et al. data show greater activation for semantic decisions on pictures relative to words in the right midfusiform gyrus (peak at $+40, -40, -20; Z = 4.7$) and the right posterior temporal cortex (peak at $54, -58, 4; Z = 3.2$). Conversely, in the left superior and middle temporal areas reported here for verbal versus nonverbal processing, the Vandenberghe et al. data show greater activation for words than pictures ($-58\ -44\ 6; Z = 4.4; -62\ -22\ 2; Z = 4.3; -56, -6, -2; Z = 3.7$). Although Vandenberghe et al. did not report these effects because they could not exclude the contribution of perceptual confounds, the left/right functional dissociation demonstrated here was, in essence, already present in their results.

Our present results are therefore fully consistent with those reported previously, but in addition, they enable us to demonstrate a dissociation between verbal and nonverbal processing even when perceptual differences have been removed.

DISCUSSION

In this study, we used functional neuroimaging to identify verbal and nonverbal conceptual processes that are common to both the visual and auditory modalities. Consistent with previous neuroimaging studies, verbal and nonverbal stimuli, whether heard or seen, activated a shared left-hemisphere network involving the posterior superior and middle temporal gyri, the inferior frontal gyrus, and the right cerebellum. In addition, consistent with neuropsychological findings, we found evidence for a left/right functional dissociation in verbal and nonverbal processing irrespective of sensory modality. Thus, we show that although the amodal conceptual system shared by verbal and nonverbal stimuli is mostly left lateralized, there is also differential activation for verbal and nonverbal stimuli. In contrast to neuropsychological case studies, whole-brain functional neuroimaging investigations enable us to localize where in the left and right hemispheres the verbal/nonverbal conceptual processing dissociation emerges. Specifically, left anterior and posterior superior temporal areas were more involved in comprehending words—heard or read—and the right midfusiform and posterior middle temporal cortices were more involved in making sense of environmental sounds and images. This dissociation was amodal (common to auditory and visual stimuli) and not affected by changes in attention or working memory demands (i.e., there were no interactions with the different conceptual tasks performed).

Previous functional imaging studies may have failed to find this double dissociation because they focused on the left-lateralized temporofrontal areas that are commonly activated by all kinds of cognitive operations (Figure 3). Indeed, the previous studies attempted to control for perceptual processing by comparing conceptual and perceptual tasks on the same set of stimuli (Vandenberghe, Price, et al., 1996). Here, in contrast, we controlled for perceptual processing by looking for differences between verbal and nonverbal stimuli independent of sensory modality (visual or auditory). As the visual and auditory modalities do not share low-level perceptual processing, any common dissociation between verbal and nonverbal processing must occur at a higher level.

Our conceptual tasks for all types of stimuli required (i) access to the long-term memory representation of meaningful objects, (ii) retrieval of their functional properties, and (iii) in the case of sequence interpretation,

evaluation of the recognized sequence in light of knowledge about objects and event successions in the real world. In addition to these shared requirements, participants may also have activated processes that were implicitly generated by the stimuli themselves. Auditory and visual verbal stimuli, for example, are more likely to evoke speech production and morphosyntactic associations than nonverbal stimuli. Indeed, the left posterior superior temporal area that was activated for verbal relative to nonverbal conditions in our experiment has previously been associated with speech production processes (Warren, Wise, & Warren, 2005; Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Okada, Smith, Humphries, & Hickok, 2003; Blank, Scott, Murphy, Warburton, & Wise, 2002; Wise et al., 2001), whereas the left anterior superior temporal cortex has been associated with morphosyntactic processing (Bornkessel, Zysset, Friederici, von Cramon, & Schlesewsky, 2005; Humphries, Love, Swinney, & Hickok, 2005; Friederici & Kotz, 2003; Friederici, Ruschemeyer, Hahne, & Fiebach, 2003; Vandenberghe, Nobre, & Price, 2002; Dronkers, Redfern, & Knight, 2000; Stowe et al., 1999; Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 1994).

With respect to the increased activation for nonverbal relative to verbal stimuli, the posterior superior temporal cortex has been associated with spatial localization in both the auditory (Rauschecker & Tian, 2000; Rauschecker, 1998a, 1998b) and visual (Milner & Goodale, 1993) domains, whereas the midfusiform is part of the ventral object recognition pathway (Haxby et al., 1991; Ungerleider, Desimone, Galkin, & Mishkin, 1984). A recent study (von Kriegstein, Kleinschmidt, Sterzer, & Giraud, 2005) has also demonstrated that familiar voice and face recognition activate the right midfusiform at Talairach coordinates [+42, −45, −24], which correspond very closely to the right fusiform activation at [+46, −46, −22] that we identified for nonverbal more than verbal conceptual processing. We therefore suggest that the increased activation in ventral and dorsal temporal regions for nonverbal stimuli reflect increased demands on the so-called ''what'' and ''where'' pathways (Haxby et al., 1991; Ungerleider et al., 1984).

The shared and distinct substrates involved in conceptual processing that we have identified provide predictions for neuropsychological studies. For example, we predict that damage to the left posterior superior and middle temporal gyri and inferior frontal regions commonly activated by verbal and nonverbal stimuli will impair both verbal and nonverbal conceptual processing (Saygin et al., 2003; Clarke et al., 2000; Vignolo, 1982). What is more, our results allow us to predict that (a) damage to the left anterior superior temporal areas that are selectively more involved in verbal than nonverbal tasks will impair verbal more than nonverbal processing (Seliger et al., 1991; Coltheart, 1980) and (b) damage to the right midfusiform and right posterior middle temporal regions that are selectively more involved in nonverbal than verbal processing will impair nonverbal more than verbal processing (Warrington, 1982; Spreen et al., 1965).

The deliberate manipulation of task difficulty— sequence interpretation versus categorization—enabled two different contexts to be probed: one in which few attentional and working memory resources are needed and one in which such orthogonal cognitive processes are involved. Crucially, although the behavioral performance was significantly different across tasks in both modalities, it was unaffected by the type of stimulus presented within each task or across sensory modality (Figure 2). This pattern of results demonstrates the successful manipulation of task difficulty in the scanner and the absence of difference in the interpretability of verbal and nonverbal stimuli in both the auditory and visual experiments.

We conclude that the functional dissociation in verbal and nonverbal conceptual processing that was previously reported in the auditory modality (Thierry et al., 2003) generalizes to the visual modality. Thus, a left versus right verbal/nonverbal functional double dissociation is observed irrespective of sensory modality, which implies that the effects originate at a conceptual, not perceptual, level. Our findings suggest that functional differences between the left and right superior temporal cortices for elementary auditory processing (e.g., von Kriegstein & Giraud, 2004; von Kriegstein, Eger, et al., 2003; Zatorre et al., 2002; Belin et al., 2000) and between the left and right inferior temporal lobes for visual processing (e.g., Cohen, Lehericy, et al., 2002; Kanwisher et al., 1997) extend to conceptual operations. In particular, we suggest that the right midfusiform gyrus may play a key role in nonverbal conceptual operations requiring high-level integration of visual and auditory information.

REFERENCES

Beauvois, M. F., & Saillant, B. (1985). Optic aphasia for colours, and colour agnosia: A distinction between visual and visuo verbal impairments in the processing of colours. *Cognitive Neuropsychology, 2,* 1–48.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature, 403,* 309–312.

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified

by functional magnetic resonance imaging. *Journal of Neuroscience, 17,* 353–362.

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

Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *Neuroimage, 26,* 221–233.

Bright, P., Moss, H., & Tyler, L. K. (2004). Unitary vs multiple semantics: PET studies of word and picture processing. *Brain and Language, 89,* 417–432.

Caramazza, A., Hillis, A. E., Rapp, B. C., & Romani, C. (1990). Multiple semantics or multiple confusions? *Cognitive Neuropsychology, 7,* 161–168.

Clarke, S., Bellmann, A., Meuli, R. A., Assal, G., & Steck, A. J. (2000). Auditory agnosia and auditory spatial deficits following left hemispheric lesions: Evidence for distinct processing pathways. *Neuropsychologia, 38,* 797–807.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123,* 291–307.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain, 125,* 1054–1069.

Coltheart, M. (1980). Deep dyslexia: A right hemisphere hypothesis. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia.* London: Routledge & Kegan Paul.

Coltheart, M., Inglis, L., Cupples, L., Mitchie, P., Bates, A., & Budd, W. (1998). A semantic subsystem of visual attributes. *Neurocase, 4,* 353–370.

Démonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., Rascol, A., & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain, 115,* 1753–1768.

Démonet, J. F., Thierry, G., & Cardebat, D. (2005). Renewal of the neurophysiology of language: Functional neuroimaging. *Physiological Reviews, 85,* 49–95.

Dronkers, N. F., Redfern, B. B., & Knight, R. T. (2000). The neural architecture of language disorder. In Michael S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 949–960). Cambridge, MA: MIT.

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (1994). A reconsideration of the brain areas involved in the disruption of morphosyntactic comprehension. *Brain and Language, 74,* 461–462.

Engelien, A., Silbersweig, D., Stern, E., Huber, W., Doring, W., Frith, C., & Frackowiak, R. S. (1995). The functional anatomy of recovery from auditory agnosia. A PET study of sound categorization in a neurological patient and normal controls. *Brain, 118,* 1395–1409.

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: Functional imaging and lesion studies. *Neuroimage, 20*(Suppl. 1), S8–S17.

Friederici, A. D., Ruschemeyer, S. A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cerebral Cortex, 13,* 170–177.

Friston, K. J., Holmes, A., Worsley, K. J., Poline, J.-B., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2,* 189–210.

Fujii, T., Fukatsu, R., Watabe, S., Ohnuma, A., Teramura, K., Kimura, I., Saso, S., & Kogure, K. (1990). Auditory sound agnosia without aphasia following a right temporal lobe lesion. *Cortex, 26,* 263–268.

Giraud, A. L., & Price, C. J. (2001). The constraints functional neuroimaging places on classical models of auditory word processing. *Journal of Cognitive Neuroscience, 13,* 754–765.

Gorno-Tempini, M. L., Price, C. J., Josephs, O., Vandenberghe, R., Cappa, S. F., Kapur, N., Frackowiak, R. S., & Tempini, M. L. (1998). The neural systems sustaining face and proper-name processing. *Brain, 121,* 2103–2118.

Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., Herscovitch, P., Schapiro, M. B., & Rapoport, S. I. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences, U.S.A., 88,* 1621–1625.

Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory–motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience, 15,* 673–682.

Humphries, C., Love, T., Swinney, D., & Hickok, G. (2005). Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Human Brain Mapping, 26,* 128–138.

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

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302–4311.

Lambon Ralph, M. A., Graham, K. S., Patterson, K., & Hodges, J. R. (1999). Is a picture worth a thousand words? Evidence from concept definitions by patients with semantic dementia. *Brain and Language, 70,* 309–335.

Levitin, D. J., & Menon, V. (2003). Musical structure is processed in ''language'' areas of the brain: A possible role for Brodmann area 47 in temporal coherence. *Neuroimage, 20,* 2142–2152.

Lichtheim, L. (1885). On aphasia. *Brain, 7,* 433–484.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience, 4,* 540–545.

McCarthy, R. A., & Warrington, E. K. (1988). Evidence for modality-specific meaning systems in the brain. *Nature, 334,* 428–430.

Milner, A. D., & Goodale, M. A. (1993). Visual pathways to perception and action. *Progress in Brain Research, 95,* 317–337.

Mummery, C. J., Ashburner, J., Scott, S. K., & Wise, R. J. (1999). Functional neuroimaging of speech perception in six normal and two aphasic subjects. *Journal of the Acoustical Society of America, 106,* 449–457.

Mummery, C. J., Patterson, K., Hodges, J. R., & Price, C. J. (1998). Functional neuroanatomy of the semantic system: Divisible by what? *Journal of Cognitive Neuroscience, 10,* 766–777.

Okada, K., Smith, K. R., Humphries, C., & Hickok, G. (2003). Word length modulates neural activity in auditory cortex during covert object naming. *NeuroReport, 14,* 2323–2326.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature, 362,* 342–345.

Price, C. J. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy, 197,* 335–359.

Rauschecker, J. P. (1998a). Cortical processing of complex sounds. *Current Opinion in Neurobiology, 8,* 516–521.

Rauschecker. J. P. (1998b). Parallel processing in the auditory cortex of primates. *Audiology and Neuro-otology, 3,* 86–103.

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of ''what'' and ''where'' in auditory cortex. *Proceedings of the National Academy of Sciences, U.S.A., 97,* 11800–11806.

Riddoch, M. J., Humphreys, G. W., Coltheart, M., & Funnel, E. (1988). Semantic systems or system? Neuropsychological evidence re-examined. *Cognitive Neuropsychology, 5,* 3–25.

Saffran, E. M., & Schwartz, M. F. (1994). Of cabbages and things: Semantic memory from a neuropsychological perspective—A tutorial review. *Attention and Performance, 25,* 507–536.

Saygin, A. P., Dick, F., Wilson, S. M., Dronkers, N. F., & Bates, E. (2003). Neural resources for processing language and environmental sounds: Evidence from aphasia. *Brain, 126,* 928–945.

Schnider, A., Benson, D. F., Alexander, D. N., & Schnider-Klaus, A. (1994). Non-verbal environmental sound recognition after unilateral hemispheric stroke. *Brain, 117,* 281–287.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123,* 2400–2406.

Seliger, G. M., Lefever, F., Lukas, R., Chen, J., Schwartz, S., Codeghini, L., & Abrams, G. (1991). Word deafness in head injury: Implications for coma assessment and rehabilitation. *Brain Injury, 5,* 53–56.

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain, 115,* 15–36.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge: Cambridge University Press.

Shallice, T. (1993). Multiple semantics: Whose confusions? *Cognitive Neuropsychology, 10,* 251–261.

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex, 6,* 11–20.

Spreen, O., Benton, A. L., & Fincham, R. W. (1965). Auditory agnosia without aphasia. *Archives of Neurology, 13,* 84–92.

Stowe, L. A., Paans, A. M., Wijers, A. A., Zwarts, F., Mulder, G., & Vaalburg, W. (1999). Sentence comprehension and word repetition: A positron emission tomography investigation. *Psychophysiology, 36,* 786–801.

Thierry, G., Giraud, A. L., & Price, C. (2003). Hemispheric dissociation in access to the human semantic system. *Neuron, 38,* 499–506.

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

Trojano, L., Grossi, D., Linden, D. E., Formisano, E., Goebel, R., Cirillo, S., Elefante, R., & Di Salle, F. (2002). Coordinate and categorical judgements in spatial imagery. An fMRI study. *Neuropsychologia, 40,* 1666–1674.

Ungerleider, L. G., Desimone, R., Galkin, T. W., & Mishkin, M. (1984). Subcortical projections of area MT in the macaque. *Journal of Comparative Neurology, 223,* 368–386.

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

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383,* 254–256.

Vignolo, L. A. (1982). Auditory agnosia. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 298,* 49–57.

von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Brain Research, Cognitive Brain Research, 17,* 48–55.

von Kriegstein, K., & Giraud, A. L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage, 22,* 948–955.

von Kriegstein, K., Kleinschmidt, A., Sterzer, P., & Giraud, A. L. (2005). Interaction of face and voice areas during speaker recognition. *Journal of Cognitive Neuroscience, 17,* 367–376.

Warren, J. E., Wise, R. J., & Warren, J. D. (2005). Sounds do-able: Auditory–motor transformations and the posterior temporal plane. *Trends in Neurosciences. 28,* 636–643.

Warrington, E. K. (1982). Neuropsychological studies of object recognition. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 298,* 15–33.

Warrington, E. K., & Crutch, S. J. (2004). A circumscribed refractory access disorder: A verbal semantic impairment sparing visual semantics. *Cognitive Neuropsychology, 21,* 299–315.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107,* 829–854.

Whatmough, C., Chertkow, H., Murtha, S., & Hanratty, K. (2002). Dissociable brain regions process object meaning and object structure during picture naming. *Neuropsychologia, 40,* 174–186.

Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., & Warburton, E. A. (2001). Separate neural subsystems within 'Wernicke's area.' *Brain, 124,* 83–95.

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex, 11,* 946–953.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences, 6,* 37–46.