What is Involved and What is Necessary for Complex Linguistic and Nonlinguistic Auditory Processing: Evidence from Functional Magnetic Resonance Imaging and Lesion Data

Frederic Dick[1,2], Ayse Pinar Saygin[2,3], Gaspare Galati[4,5], Sabrina Pitzalis[5,6], Simone Bentrovato[6], Simona D'Amico[6], Stephen Wilson[2,7], Elizabeth Bates[2], and Luigi Pizzamiglio[5,6]

Abstract

■ We used functional magnetic resonance imaging (fMRI) in conjunction with a voxel-based approach to lesion symptom mapping to quantitatively evaluate the similarities and differences between brain areas involved in language and environmental sound comprehension. In general, we found that language and environmental sounds recruit highly overlapping cortical regions, with cross-domain differences being graded rather than absolute. Within language-based regions of interest, we found that in the left hemisphere, language and environmental sound stimuli evoked very similar volumes of activation, whereas in the right hemisphere, there was greater activation for environmental sound stimuli. Finally, lesion symptom maps of aphasic patients based on environmental sounds or linguistic deficits [Saygin, A. P., Dick, F., Wilson, S. W., Dronkers, N. F., & Bates, E. Shared neural resources for processing language and environmental sounds: Evidence from aphasia. *Brain, 126,* 928–945, 2003] were generally predictive of the extent of blood oxygenation level dependent fMRI activation across these regions for sounds and linguistic stimuli in young healthy subjects. ■

INTRODUCTION

There has been a rekindling of interest in exploring the links between putatively symbolic cognitive skills, such as language and mathematics, and those underlying perception and movement (Gentilucci, 2003; Bates & Dick, 2002). More specifically, "embodied" theories of cognition (MacWhinney, 1999) suggest that higher level linguistic processing may be enmeshed within the perceptual and motor substrates that allow us to hear and produce complex, meaningful sounds. If this view is correct, then language might share developmental trajectories and processing resources with nonlinguistic skills that have similar sensorimotor demands. For example, neuroimaging research on music perception has suggested a substantial degree of overlap between brain regions underlying music perception and language comprehension (for review and discussion, see Koelsch & Siebel, 2005). But although music perception entails many processing demands similar to language compre-

hension (e.g., acoustical, syntactic, memory), music does not share one of language's most important properties in that it does not typically involve reference—the indexing of one object or event with a (possibly arbitrary) sound-based sign or symbol.

Unlike most music, environmental sounds can have an iconic or indexical relationship with the source of the sound and thus can be a useful method for comparing meaningful linguistic and nonlinguistic auditory comprehension. Environmental sounds can be defined as sounds generated by real events that gain sense or meaning by their association with those events (Ballas & Howard, 1987). Spoken words and environmental sounds share many spectral and temporal characteristics, and recognition of both classes of sounds breaks down in similar ways under acoustical degradation (Gygi, Kidd, & Watson, 2004). Studies with adult subjects have shown that, like words, processing of individual environmental sounds is modulated by contextual cues (Ballas & Howard, 1987) and item familiarity and frequency (Cycowicz & Friedman, 1998; Ballas, 1993). Environmental sounds can prime semantically related words (Van Petten & Rheinfelder, 1995) and may prime other semantically related sounds (Stuart & Jones, 1995; but cf. Friedman, Cycowicz, & Dziobek, 2003; Chiu &

[1]Birkbeck College, London, UK, [2]University of California, San Diego, [3]University College London, UK, [4]University G. d'Annunzio, Italy, [5]Fondazione Santa Lucia IRCSS, Italy, [6]Università La Sapienza, Italy, [7]University of California, Los Angeles

Schacter, 1995). Spoken language and environmental sound comprehension show similar developmental trajectories in infants and toddlers (Cummings, Saygin, Bates, & Dick, 2005), typically developing school-age children, and children with language impairment and perinatal focal lesions (Borovsky et al., 2006).

Environmental sounds also differ from speech in several ways. Unlike spoken words, individual environmental sounds are causally bound to the sound source or referent, unlike the semiarbitrary linkage between a word's pronunciation and its referent. The "lexicon" of environmental sounds is small and semantically stereotyped and clumpy; neither are these sounds easily recombined into novel sound phrases (Ballas, 1993). There is quite wide individual variation in exposure to different sounds (Cummings, Saygin, et al., 2005; Gygi et al., 2004), and, correspondingly, healthy adults show considerable variability in their ability to recognize and identify these sounds (Saygin, Dick, & Bates, 2005). Finally, most environmental sounds are not produced by the human vocal tract. In fact, the neural mechanisms of nonlinguistic environmental sounds that can and cannot be produced by the human body appear to differ significantly (Lewis, Brefczynski, Phinney, Janik, & DeYoe, 2005; Pizzamiglio et al., 2005; Aziz-Zadeh, Iacoboni, Zaidel, Wilson, & Mazziotta, 2004).

Despite these differences, there is evidence suggesting that comprehension of environmental sounds and spoken language recruit similar mechanisms when task and stimulus demands are well matched (reviewed in Saygin et al., 2005; Saygin, Dick, Wilson, Dronkers, & Bates, 2003). For instance, when presented in semantically matching or mismatching contexts, the electrophysiological response to environmental sounds and their spoken labels is quite similar in distribution and latency, as demonstrated in school-age children (Cummings, Dick, Saygin, Townsend, & Ceponiene, 2005) and in healthy adults (Cummings et al., 2006; Van Petten & Rheinfelder, 1995, but cf. Plante, Van Petten, & Senkfor, 2000; see Discussion). In studies of patients with brain injury and aphasia, Varney (1980) and Schnider, Benson, Alexander, and Schnider-Klaus (1994) found that deficits in language processing were comorbid with deficits in environmental sound recognition (but cf. Clarke, Bellmann, Meuli, Assal, & Steck, 2000; Clarke, Bellmann, De Ribaupierre, & Assal, 1996, who indirectly compared cross-domain performance). Using closely matched linguistic and nonlinguistic stimuli and the same task across both domains, Saygin et al. (2003) showed that deficits in the two domains were tightly correlated over 29 patients with left-hemisphere damage. Furthermore, damage to posterior left middle and superior temporal gyri and to the inferior parietal lobe was the best predictor of deficits in processing for both environmental sounds and spoken language. Surprisingly, classical "Wernicke's area" lesions were more detrimental for processing nonverbal sounds than for verbal sounds.

In healthy subjects, direct comparisons between language and environmental sound comprehension have been made using functional magnetic resonance imaging (fMRI; Specht & Reul, 2003; Humphries, Willard, Buchsbaum, & Hickok, 2001) and positron emission tomography (PET; Thierry, Giraud, & Price, 2003; Giraud & Price, 2001). ("Lower level" speech perception has also been contrasted with environmental sounds by using PET; Belin, Zatorre, Lafaille, Ahad, & Pike, 2000.) Given the tight cross-domain correlation found in the neuropsychological lesion-mapping study of Saygin et al. (2003), one might expect to see in these studies a similarly high degree of overlap in fMRI and PET activation, at least within the "dominant" left hemisphere (LH). Indeed, both fMRI studies showed language- and environmental-sound-related activation in regions including the bilateral transverse temporal gyri, precentral and inferior frontal gyri, and superior and middle temporal gyri. Conjunction analyses in the PET studies showed significant cross-domain overlap in bilateral middle and superior temporal gyri, the left inferior frontal gyrus (IFG), and the right and left cerebellar hemispheres. All the above imaging studies also showed some differences in activation for language and environmental sounds in both hemispheres. However, the location and extent of these differences diverged considerably over studies.

For instance, whereas Humphries et al. (2001) showed a greater bilateral anterior temporal lobe activation for language than environmental sounds, this was not the case in Specht and Reul (2003), who found no differences between conditions in this region. In addition, Humphries et al. showed a right IFG advantage for environmental sounds; there was no indication of this difference in Specht and Reul. Moreover, in their volume-of-interest analyses, Specht and Reul showed a strong *sounds > language* advantage in the right and left planum temporale and transverse gyri, whereas the results of Humphries et al. showed no differences for these regions in the right hemisphere, but with indications of the *converse* effect (language > sounds) in the left planum temporale. Finally, Thierry et al. (2003) showed more language- than environmental-sound-related PET activation in patches along the left superior temporal gyrus (STG), descending into the sulcus, as well as in the left cerebellum; in contrast, environmental sound stimuli elicited more activation in the right posterior STG than did words—an effect that was only observed with an "active" task used by Thierry et al., and not with the "passive" task used in the companion study by Giraud and Price (2001).

These cross-study differences are puzzling and may depend in part upon the exact stimuli used, methodological details, and task constraints. For instance, even subtle differences in semantic content between linguistic and nonlinguistic stimuli may drive differential profiles of activation (for an example of this in the purely

language domain, see Tettamanti et al., 2005; Hauk, Johnsrude, & Pulvermuller, 2004). Auditory masking from MRI scanner noise (Hall et al., 1999) may also exert differential effects across domains and across stimuli sets. Finally, the power to detect cross-domain similarities and differences, especially across hemispheres, may be substantially modulated by variations in data acquisition, anatomical morphing, image processing, and statistical techniques (Saad, Ropella, DeYoe, & Bandettini, 2003; Lazar, Luna, Sweeney, & Eddy, 2002; Fischl, Sereno, Tootell, & Dale, 1999). In particular, the intersubject morphological variability around the sylvian fissure can introduce substantial error in fMRI spatial normalization, to the extent that activation can ''jump'' over sulci (Ozcan, Baumgartner, Vucurevic, Stoeter, & Treede, 2005).

The relative disparity between the findings of Saygin et al. (2003) (where linguistic and nonlinguistic deficits and their left-hemisphere lesion correlates were tightly yoked) and those of some of the functional imaging studies (showing substantial cross-domain differences in left- and right-hemisphere activation) raises the larger question of the relationship between the neural substrates that are *necessary* for processing (as revealed by lesion analysis) versus those that are *involved* in processing (as assessed by neuroimaging; for discussion, see Casey, Tottenham, & Fossella, 2002; Moses & Stiles, 2002). Given the sometimes divergent results from lesion and functional activation comparisons in other domains (e.g., with face perception [Bouvier & Engel, 2005] and music [Koelsch & Siebel, 2005]), it may be that these two neuropsychological methods are tapping into different aspects of neural function. For instance, fMRI activation likely reflects neural activity in the cortical mantle and subcortical nuclei, whereas lesion mapping results may reflect contributions from both gray matter and underlying white matter tracts. (For an interesting discussion of the neurophysiological reasons that might underlie differences between functional and structural measures of behavioral deficits, see Dronkers & Ogar, 2004; Hillis et al., 2004.)

We performed the present fMRI comparison of environmental sound and language comprehension to accomplish three basic aims. The first was to untangle the thicket of contradictory neuroimaging findings discussed above by explicitly addressing the semantic, acoustical, and attentional factors that differ over these studies. In order to balance semantic and/or conceptual/perceptual information over environmental sound and language stimuli, we used the same stimulus norming procedure as reported in Saygin et al. (2003, 2005). We compared naturalistic environmental sounds matched to their empirically derived linguistic counterparts, with sounds representing a wide range of semantic categories and varying considerably in duration—as do sounds in the ''acoustical wild.'' In order to minimize the influence of scanner noise while maximizing statistical power, we

employed a ''sparse sampling'' image-acquisition protocol (Lewis et al., 2004; Hall et al., 1999) with a blocked design, using a single control task to preclude subtle task-related confounds. To assure that subjects were performing similarly in both conditions and to facilitate cross-methodology comparisons, we used the same sound–picture matching task as was used in Saygin et al. (2003). Here, subjects listened to either an environmental sound or a short phrase and at the same time saw two black-and-white line drawings, one of which was closely related to the sound. They then pushed a mouse button to indicate which drawing best corresponded to the sound or phrase. This task was originally adapted from clinical tests into an ''online'' measure to maintain continuity with most of the early neuropsychological work in this field (see Saygin et al., 2003, for a review).

Our second aim was to directly assess the relative lateralization of environmental sound- and language-related activation in perisylvian regions. Thus, we conducted not only a standard whole-brain group analysis of intensity of activation for comparison with previous studies, but also regions-of-interest (ROI) analysis on the relative volume and intensity of activation in the bilateral perisylvian cortex, one that explicitly takes into account variations in individual subject's cortical anatomy and that allows for a statistically powerful test of volume of activation in a well-defined region of cortex directly comparable over hemispheres.

Our final aim was to quantitatively compare neuroimaging and neuropsychological brain mapping data. Here, we used a variant of a new lesion analysis method (voxel-based lesion symptom mapping [VLSM]; Bates et al., 2003) to quantify the relationship between lesion maps derived from a complementary study of left-hemisphere-injured patients' environmental sound and language comprehension (previously reported in Saygin et al., 2003) and the activation maps derived from the present study.

METHODS

Imaging Protocol

Fifteen young native speakers of Italian with no known neurological abnormalities (7 women and 8 men; age, 22–33 years) were scanned on a 1.5-T Siemens (Erlangen, Germany) Vision clinical scanner equipped with a standard head coil at the Fondazione Santa Lucia, Rome, Italy. Using a low-bandwidth EPI sequence (TR = 11, TE = 50, flip angle = 90°, 64 × 64 matrix, FOV = 192), we acquired four runs of functional data (7′36″, 40 volumes total). Thirty axial slices were collected sequentially with an in- and through-plane resolution of 3 mm (1-mm gap between slices). A sparse sampling sequence (Hall et al., 1999) was used where image acquisition occurred only during the first 3 sec of the

11-sec TR, thus allowing the hemodynamic response to the acoustical noise produced by the gradient coils to return almost to baseline before the next acquisition. Experimental stimulation began after four TRs to allow for field stabilization. After functional scanning we acquired a single high-resolution structural volume by using a magnetization-prepared rapid gradient-echo (MPRAGE) sequence (TR = 11.4 msec, TE = 4.4 msec, flip angle = 100°, voxel = 1 × 1 × 1 mm, 220 coronal slices).

The PsyScope experimental driver was used to present stimuli and collect response data (Cohen, MacWhinney, Flatt, & Provost, 1993). Visual stimuli were projected from a specially configured video projector onto a custom-built screen fitted over the head coil. Auditory stimuli were delivered through pneumatic headphones fitted with high-quality sound defenders; a specially installed speaker also delivered auditory stimuli into the magnet suite. Button-press data were collected by using an opto-isolated two-button response box (Stark Labs, San Diego, CA).

Experimental Design and Task

We used a blocked design in order to maximize statistical power for single-subject analyses, where a single run consisted of twelve 33-sec blocks alternating between experimental and control conditions. Two runs alternated between environmental sounds and control blocks, and two between language and control blocks, with run order counterbalanced across subjects. There were eight trials per block, with a 1750-msec intertrial interval; the beginning of each block was synchronized with the onset of a TR.

An experimental condition trial consisted of simultaneous presentations of an environmental sound or linguistic description and two black-and-white line drawings presented in white frames on the left and right sides of the video display (see Figure 1 for schematic); the participant selected the picture that best matched the sound by pressing the left or right mouse button as quickly and accurately as possible.

The control condition matched the experimental task for basic auditory, visual, motor, and attentional demands, and was the same for both environmental sound and language conditions. The control condition differed from the experimental condition only in that the sound stimulus consisted of two identical or distinct simple tones presented sequentially, whereas visual stimuli consisted of two white frames containing two nonsense shapes, one frame containing two identical shapes, and the other containing two distinct shapes. The participant pushed the left or right mouse button to indicate which nonsense-shape pair matched the tones (e.g., if the tones differed, then the participant chose the frame with the different shapes, whereas if the tones were identical, he or she chose the frame with identical shapes).

Stimuli

In order to ensure that our environmental sound stimuli and linguistic equivalents were culturally and linguistically consistent, we performed a new norming study with native Italian speakers at the Università La Sapienza in Rome, using the same methodology as the English-language version used for the Saygin et al. (2003) study and described at length in Saygin et al. (2005). Full details of the norming experiment can be found at crl.ucsd.edu/experiments/envsoundsfMRI. Briefly, 96 environmental sounds were selected by using the new Italian norms; the linguistic equivalents to these sounds

Figure 1. Schematic of (A) experimental and (B) control tasks. In the control task, the subject pressed the button under the picture matching the sound or linguistic description (in this case, the subject should pick the right picture, the bus). In the control condition, the subject heard two beeps that were either identical or differ in frequency (here indicated by the typeface of "beep"). If the subject heard two different beeps, as in this example, he or she pushed the button under the frame containing the two different nonsense shapes (the left frame); if the beeps were identical, the frame containing two identical shapes was picked.
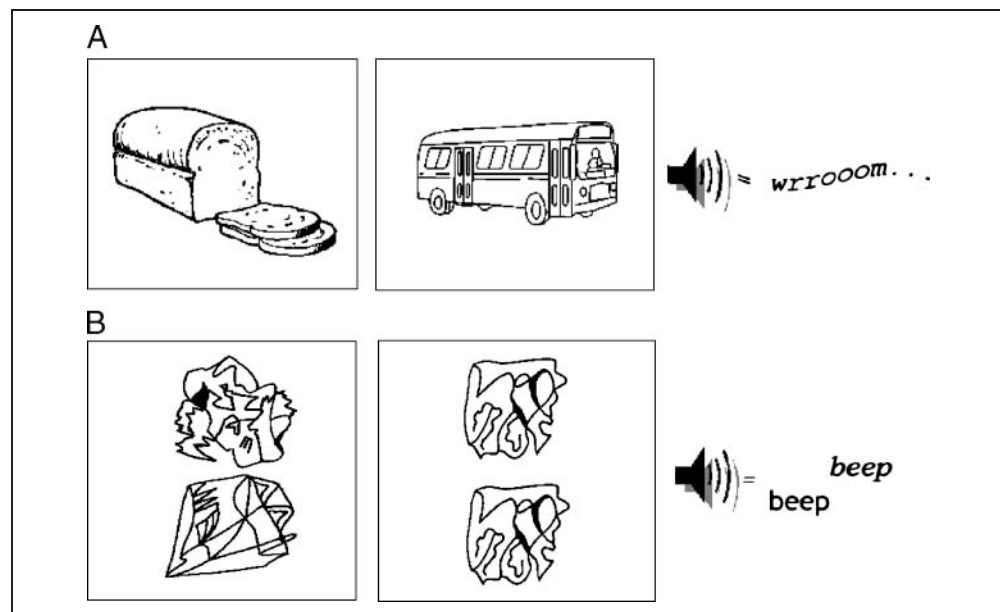
were drawn from the most common verbal phrases produced by subjects for each sound. These phrases were often of the form "[Noun] that [Verbs]," such as "Bambino che piange" (English gloss, "A baby who is crying"), but could also be simple descriptions, like "Campanella della scuola" ("school bell"). Linguistic phrases were recorded by a male Roman adult; both environmental sounds and linguistic equivalents were sampled at 44.1 kHz with 16-bit quantization. Environmental sounds (Env) and language (Lang) stimuli were overall equated for intensity to within 1 dB, but with Env stimuli necessarily varying somewhat more due to the different sound sources (Env average intensity = 73 dB [SD = 5.3], Lang = 72 dB [SD = 2.0]). Language descriptions are generally more "informationally compact" than the environmental sounds they describe and thus are on average 590 msec shorter. Unlike other stimulus characteristics like intensity, for instance, such relatively small differences in duration do not appear to have a significant effect on the blood oxygenation level dependent response to auditory stimuli (Jäncke et al., 1999).

Control stimuli were four pairs of sequential, binaurally presented 600-msec sine-wave tones (440 Hz [A1]/587.3 Hz [D2]) that either contained same or different tones; the four permutations were A1/A1, D2/D2, A1/D2, and D2/A1, with 24 identical exemplars per pair. Tone pairs were separated by 500-msec silence with a 50-msec amplitude ramp at the beginning and end of each tone. Experimental visual stimuli were black-and-white line drawings of the sound-associated objects overlaid on a white frame. Control visual stimuli were two black-and-white drawings of nonrepresentational but distinct "nonsense shapes" (Saccuman et al., 2002) displayed on a single frame; the two drawings were either identical or easily distinguished as being different.

Each environmental sound was paired with a matched or "target" drawing and a semantically unrelated "foil" (as measured by latent semantic analysis; Landauer, Foltz, & Lahan, 1998; see Saygin et al., 2005, for details). Each of the 96 environmental sounds and their corresponding linguistic descriptions was presented only once during the experiment, whereas each of the 96 line drawings was presented twice, once as target and once as foil. Control tone pairs were matched to a "target" nonsense-shape pair presented in a single frame. Same-tone pairs were matched to a nonsense-shape pair containing identical shapes, and different-tone pairs were assigned a target nonsense-shape pair containing two different shapes. Each of the 96 nonsense-shape pairs was presented twice, once as target and once as foil.

Image Processing and Analysis

Data from 3 of the 15 scanned participants were not used, with two excluded due to large-scale movement throughout scanning, and one excluded because of a neurological abnormality discovered during the scanning session. All functional data were re-registered to correct for small head movements (Friston et al., 1995) with correction for slice timing. The resulting data set was manually coregistered to the T1-weighted high-resolution volumetric image.

For whole-brain analyses, SPM99 was used to resample and spatially normalize functional images to the standard MNI template (Friston et al., 1995); voxel size after normalization was 3 × 3 × 3 mm. Data were then analyzed using a two-stage random-effects analysis (Friston, Holmes, & Worsley, 1999). First, each participant's hemodynamic response was characterized using a boxcar function convolved with a synthetic hemodynamic response function (HRF); their temporal derivatives, a constant term, a set of cosine basis functions serving as a high-pass filter, and the head movement parameters estimated during the preprocessing stage were also included in the statistical model. For each subject-specific model, linear contrasts were derived from the regression parameters; these subject-specific effect-size images were spatially smoothed using an isotropic Gaussian kernel (6-mm full width at half maximum) and then entered at the second stage into one-sample t tests. For each effect of interest, t maps were initially thresholded at a voxelwise $p < .01$ (Friston, Holmes, Poline, Price, & Frith, 1996). The significance of each cluster was then estimated by using distribution approximations from the theory of Gaussian fields, resulting in a corrected p value (Worsley, Marrett, Neelin, Friston, & Evans, 1995). Activation clusters were retained as significant at $p < .05$ corrected for whole-brain volume. Activations were displayed on cortical surface reconstructions (FreeSurfer; Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999; Fischl, Sereno, Tootell, et al., 1999) and anatomically labeled by the BrainShow software, using a parcellation of the MNI single-subject brain (Tzourio-Mazoyer et al., 2002).

For ROI analyses, we used FreeSurfer to generate a full complement of cortical surfaces for each subject. We then used a novel computational technique included in FreeSurfer that probabilistically incorporates geometrical and neuroanatomical information to parcellate cortex with an accuracy equal or superior to that of highly reliable manual methods (Fischl et al., 2004; borders and labeling conventions follow those of Duvernoy, 1999). A subset of parcellated ROIs was selected (see the Introduction and Results) and projected into a copy of the subject's manually aligned EPI volume in native space. The set of ROIs included a three-part subparcellation of the superior temporal sulcus (STS) and middle temporal gyrus (MTG) (running anterior to posterior) that is not included in the FreeSurfer atlas. The complete set of ROIs (including detailed anatomical descriptions and additional technical details) is available at crl.ucsd.edu/experiments/envsoundsfMRI. (Figure 3 shows each ROI

projected out to the pial surface in a single subject, temporal lobe subparcellations not shown in the figure.)

ROI statistical analyses were performed by using AFNI software (Cox, 1996). After applying a 6-mm isotropic spatial blur within a ''brain-only'' mask to each EPI run, we concatenated the two runs from each condition and then performed voxelwise simultaneous multiple linear regression. As in the whole-brain analyses, the model parameters included the HRF-convolved boxcar waveform, separate DC, linear, and quadratic trends for each run, as well as the head movement estimates. (These AFNI-generated individual t maps are very similar to those generated in the first level of SPM analysis but are not warped to a common stereotactic template.) For all voxel counts, we used an ROI-wise false discovery rate (FDR)-corrected p value of .05 (Benjamini & Hochberg, 1995), including positively signed voxels only and discarding negatively signed ones.

Overlap within ROIs was calculated as the proportion of voxels active in both conditions divided by voxels active in the least active condition. In other words, this proportion represents the percentage of voxels from the smaller of the two activations that fall within the volume occupied by the larger one. We compared degree of overlap at strict (ROI-wise $p < .05$) and lenient (voxelwise $p < .05$) statistical thresholds. The FDR-based proportions represent data from approximately half the subjects (e.g., not every subject had suprathreshold activation in each ROI), whereas the voxelwise-thresholded data are more representative of the entire sample of subjects.

Activation–Lesion Correlation Analyses

Lesion maps were morphed versions of those reported in the companion neuropsychological study (Saygin et al., 2003). In that study, 30 left-hemisphere-damaged (LHD) patients and 5 right-hemisphere-damaged (RHD) patients were tested behaviorally using the same paradigm and a subset of the stimuli used in the present experiment. All patients were right-handed, native English speakers. Patients' computed tomography (CT) or MRI scans and medical records were evaluated by a licensed neurologist; only patients who had unilateral lesions due to a single cerebrovascular accident and who did not have diagnosed or suspected hearing or (uncorrected) visual difficulties, dementia, head trauma, tumors, or multiple infarcts were included. Data from one patient was excluded due to the possibility of a second infarct.

For 20 of the LHD patients, computerized lesion reconstructions to be used in lesion overlay analyses were available. Lesion reconstructions were available only for two of the RHD patients who participated in this study so were not included in our lesion analyses.

Lesion reconstructions were based on CT or MRI scans at least 3 weeks after onset (most scans were performed several months after the stroke) and were hand-drawn onto 11-axial-slice templates based on the atlas of DeArmond, Fusco, and Dewey (1976), then entered into a Macintosh computer via electronic bitpad using software developed at the VA Medical Center in Martinez, California (Frey, Woods, Knight, Scabini, & Clayworth, 1987). All reconstructions were completed by the same board-certified neurologist experienced in neuroradiology but blind to the behavioral deficits of the patients. The reliability of these lesion reconstructions has been verified (see Knight, Scabini, Woods, & Clayworth, 1988) and similar techniques have been used by many laboratories using different templates (e.g., Bouvier & Engel, 2006; Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000).

Using VLSM software (Bates et al., 2003), we constructed maps of brain areas associated with deficits by making overlays of patients who show substantial deficits in either environmental sound or speech processing, and identifying the regions that are most commonly included in a number of these patients. For this purpose, patients whose accuracy scores were two or more standard deviations lower than the average score for age-matched controls were considered ''impaired'' in either environmental sounds ($n = 8$) or language performance ($n = 10$).

The lesion overlays were morphed into EPI masks for the current study as follows. A T1 (MPRAGE) whole-brain volume from the current study was visually chosen as being most similar to the brain used in the atlas upon which the 11-axial-slice VLSM brain template is based (DeArmond et al., 1976). The individual subject's T1 volume was manually rotated and translated into alignment with the template volume; this T1 volume was then used to establish a linear morph into standard stereotactic (Talairach) space using AFNI. Lesion masks for functional analyses were produced by overlaying each lesion template on the aligned T1 volume, warping to Talairach space using AFNI, then resampling to 3-mm isotropic voxels (nearest-neighbor interpolation). Although there was significant and important interindividual anatomical variability, the general accuracy of the morph was confirmed by sending points back and forth between anatomical landmarks on the atlas and the T1 volume of each subject when both were morphed into Talairach space.

For activation–lesion correlation, we used the same activation maps generated for the ROI analyses (see above). To correct for multiple comparisons, we calculated the FDR over all voxels contained in the lesion map; due to the spatial contingencies caused by warping and resampling, we used a correction to the FDR algorithm that allows for an arbitrary (nonindependent) distribution of p values (Genovese, Lazar, & Nichols, 2002).

RESULTS

Behavioral Results

Button-press data were analyzed for 9 of 12 subjects (3 data points were lost due to equipment failure). For

both experimental and control tasks, accuracy was at ceiling (98.5–100% correct); there were significant ($p <$ .05) cross-domain differences in reaction times (RTs) with shorter mean RTs for environmental sounds (1090 msec, $SE = 45$) than linguistic equivalents (1387 msec, $SE = 73$). These mean RTs were extremely similar to those reported for previous studies using similar stimuli (Saygin et al., 2003, 2005; Dick, Bussiere, & Saygin, 2002).[1] Control task RTs were overall slower, and were not affected by experimental condition (mean control RTs in environmental sounds blocks 2318 ($SE = 31$) msec; in language blocks 2328 ($SE = 36$) msec).

Imaging Results

Whole-brain Group Analyses

Relative to the control condition, both environmental sounds (Env) and language (Lang) evoked significant bilateral activation in the inferior frontal gyri, superior temporal gyri (anterior/transverse/posterior), and posterior middle and inferior temporal and fusiform gyri (see Table 1). Language stimuli evoked more activation than environmental sounds in only three brain regions: (1) approximately the middle third of the left hemisphere (LH) MTG, with a much smaller patch superiorly on the left STG, slightly posterior to the transverse gyrus; (2) the anterior portion of the left and right STG; and (3) lateral fusiform gyrus bilaterally onto the inferior temporal gyrus (ITG), with a much larger Lang > Env difference in the right hemisphere (RH) (see Figure 2 and Table 2A).

Environmental sounds evoked more activation than did language stimuli in two general regions (Table 2B and Figure 2): (1) the right STG, with patches scattered along the extent of the planum temporale and an additional patch in anterior and superiormost extent of the right supramarginal gyrus and (2) two patches in the right IFG.

ROI Analyses

In each ROI, we performed a Domain (Env/Lang) × Hemisphere (LH/RH) analysis of variance on the volume of FDR-corrected suprathreshold activation (see Methods). When we observed a significant main effect or interaction, we performed planned Bonferroni-corrected ($p <$ .05) a priori linear contrasts to clarify effects of lateralization and domain.

We first asked whether activation *across* domain was preferentially lateralized to one hemisphere in one or more ROIs (see Figure 3A for main ROIs, Figure 3B for subparcellated temporal lobe ROIs). The only ROI with a significant ($p <$ .05) main effect of hemisphere and no Hemisphere × Domain interaction was the planum polare, where both language and environmental sound activation was L > R (although in individual pairwise

contrasts, only language activation was significantly so). Second, we asked whether there were any ROIs where one domain showed more activation generally across both hemispheres with no significant Hemisphere × Domain interaction. Here, environmental sound activation was overall greater than for language in the IFG, opercular part, the supramarginal and angular gyri, and the subparcellated posterior STG/STS/MTG (see Figure 3A and B). Contrasts showed significant Env > Lang differences after Bonferroni correction in the RH for all four ROIs; in the LH, pairwise differences were significant only in the posterior STG/STS/MTG, with a trend in the supramarginal gyrus (.07) and nonsignificant in the other two ROIs.

We then asked whether there were cortical regions where activation was differentially modulated by domain in each hemisphere, as reflected by a Hemisphere × Domain interaction. These interactions took two general forms. In the first, language and environmental sounds evoked approximately the same degree of activation in the LH (e.g., no significant difference in pairwise comparisons of LH Lang vs. LH Env), but in the RH, environmental sounds evoked significantly more activation than did language stimuli. We observed significant interactions of this type in the IFG, orbital part, the STS, and the STG, planum temporale; a marginally significant interaction showing the same pattern was found in the IFG, triangular part ($p =$ .09).

The other type of Hemisphere × Domain interaction we observed showed the converse effect, where in the LH language evoked significantly more activation than environmental sounds, whereas in the RH there were no significant differences across domain. Significant interactions following this pattern were found in the lateral aspect of the STG, and the subparcellated middle STS/STG/MTG; the subparcellated anterior STS/STG/MTG showed the same trend but at a very marginal ($p =$ .11) level of interaction. (Note that the anterior and middle subparcellated regions included the lateral aspect of the STG.)

Finally, to obtain a metric of functional lateralization similar to that reported in previous studies, in each ROI we performed Bonferroni-corrected linear contrasts within domain and across hemispheres. As would be expected, language-evoked activation was significantly L > R in all ROIs *except* the angular and supramarginal gyri, the transverse gyri, and the opercular part of the IFG. Unlike language, environmental sounds showed significant L > R activation *only* in the ITG. However, environmental sounds did not show any significant lateralization effects in the opposite (R > L) pattern—the only ROI showing a significant trend in this direction was the planum temporale ($p =$ .084).

In order to assure ourselves that the results of these volume-of-activation contrasts were robust under different voxelwise thresholds, we reran all contrasts using an absolute voxelwise threshold of $p <$ .01. We observed no

Table 1. Location of Activation Clusters and Peak Coordinates for Main Effects of Language versus Control and Environmental Sounds versus Control

Activation Cluster	Extent (mm³)	Activation Peaks	x	y	z	Z Score
A. Language vs. control						
Left frontal	11,313	Left IFG (triangular)	−56	22	26	4.51
		Left IFG (triangular)	−46	28	18	4.12
		Left IFG (triangular)	−44	18	26	3.86
		Left middle frontal gyrus (posterior)	−38	12	36	3.50
Right frontal	1,404	Right IFG (triangular)	56	34	0	3.62
		Right IFG (triangular)	38	30	12	3.41
Left temporal	50,787	Left fusiform gyrus	−28	−48	−10	4.82
		Left middle occipital gyrus	−40	−74	14	4.24
		Left STG (anterior)	−56	−12	6	4.20
		Left MTG (posterior)	−40	−60	12	4.14
		Left ITG (posterior)	−52	−60	−12	4.01
		Left MTG (anterior)	−46	−32	0	3.94
Right temporal	40,095	Right STG (anterior)	56	−12	−4	4.91
		Right ITG (posterior)	46	−48	−16	4.76
		Right STG (anterior)	52	−18	−4	4.66
		Right STG (anterior)	58	−6	−6	4.46
		Right MTG (posterior)	46	−72	20	4.22
		Right fusiform gyrus	38	−26	−18	4.14
		Right MTG (anterior)	64	−32	0	3.88
B. Environmental sound vs. control						
Left frontal	3,429	Left IFG (triangular)	−46	30	12	4.33
		Left IFG (orbital)	−46	40	−12	3.60
		Left IFG (orbital)	−40	42	−12	3.33
Right frontal	5,157	Right IFG (opercular)	44	12	30	5.09
		Right IFG (triangular)	52	40	8	3.73
		Right IFG (triangular)	56	36	6	3.31
		Right IFG (orbital)	50	42	−6	3.11
Left temporal	33,750	Left STG (posterior)	−50	−44	12	4.28
		Left ITG (posterior)	−40	−44	−16	4.24
		Left fusiform gyrus	−38	−42	−18	4.12
		Left inferior occipital gyrus	−46	−60	−12	4.09
		Left ITG (posterior)	−46	−68	−10	4.06
		Left fusiform gyrus	−38	−26	−22	3.80
		Left MTG (posterior)	−44	−72	18	3.79
		Left STG (anterior)	−44	−14	−4	3.64
		Left STG (polar)	−28	6	−28	3.60

Table 1. (*continued*)

Activation Cluster	Extent (mm³)	Activation Peaks	x	y	z	Z Score
		Left STG (posterior)	−52	−26	8	3.57
		Left MTG (anterior)	−58	−6	−12	3.35
		Left amygdala	−28	−2	−18	3.21
Right temporal	26,487	Right STG (anterior)	62	−6	0	4.08
		Right STG (posterior)	50	−36	12	3.82
		Right STG (posterior)	40	−36	8	3.59
		Right fusiform gyrus	38	−62	−18	3.65
		Right STG (polar)	56	6	−16	3.55
		Right ITG (posterior)	50	−66	−4	3.52

Some anatomical regions may be listed more than once if there are multiple activation peaks within that region. IFG = inferior frontal gyrus; STG = superior temporal gyrus; ITG = inferior temporal gyrus.

changes in the direction, magnitude, or significance of the effects resulting from this change in threshold.

Within-ROI Overlap

Particularly in the left or ''language-dominant'' hemisphere we see very similar volumes of activation for both language and environmental sounds. However, because these are relatively large ROIs (containing between 62 and 560 functional voxels on average, with volumes of 2229–20,133 mm³), it is possible that a good portion of each ROI could be active during each task, but that these

voxels be located in completely nonoverlapping parts of the ROI.

Therefore, we asked whether the specific voxels that were active in environmental sound processing were also active in language processing and vice versa. In Figure 4, we depict the relative overlap in activation over ROIs. This is expressed as a proportion: the percentage of voxels from the smaller of the two activations that falls within the volume occupied by the larger one (see Methods for details).

In general, the majority of the smaller volume of activation was nested within the larger activation volume,
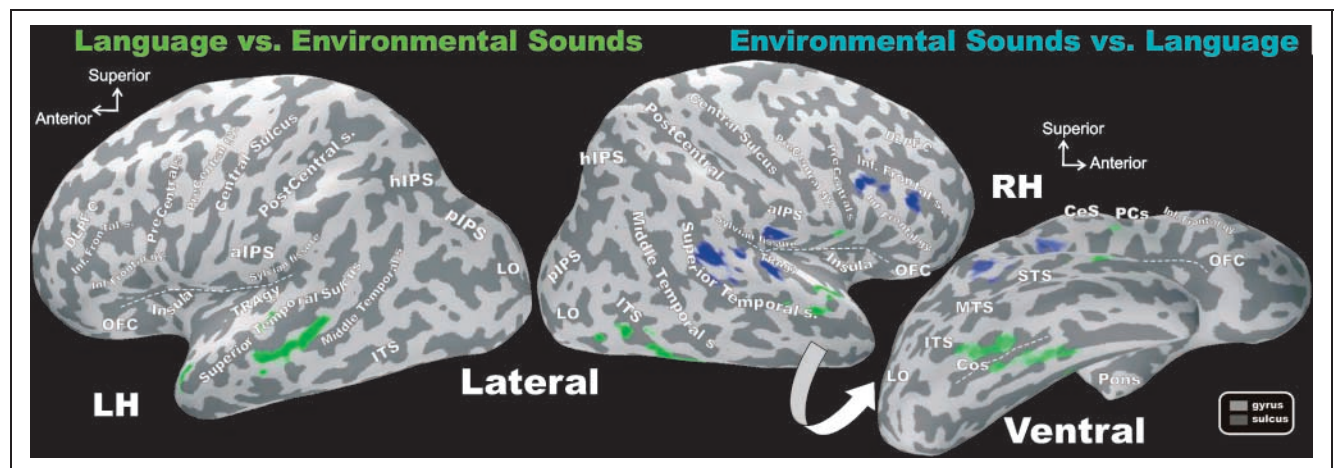

Figure 2. Significant differences in group activation between environmental sounds and language painted on the inflated left and right hemispheres of the MNI template. Green patches show regions with significantly more activation for language than for environmental sounds; blue patches show regions where activation for environmental sounds is greater than language. The main anatomical sulci (dark gray) and gyri (light gray) have text labels: OFC = orbitofrontal cortex; DLPFC = dorsolateral prefrontal cortex; PCs = precentral sulcus; CeS = central sulcus; IPS = intraparietal sulcus; aIPS, pIPS, hIPS = anterior, posterior, and horizontal segments of the intraparietal sulcus, respectively; LO = lateral occipital sulcus; TRAgy = transverse gyrus; STS = superior temporal sulcus; MTS = middle temporal sulcus; ITS = inferior temporal sulcus; CoS = collateral sulcus. Both hemispheres are shown in lateral view and only the right hemisphere is shown also in ventral view to indicate the patches in the fusiform gyrus ventrally located around the collateral sulcus where language-related activation was greater than for environmental sounds. The fundus of both the CoS and sulcal cortex from the sylvian fissure to the insula are indicated by the dashed white line. Logo next to each surface indicates two main orthogonal directions (anterior–posterior and superior–inferior).

Table 2. Location of Activation Clusters and Peak Coordinates for Language versus Environmental Sounds and Environmental Sounds versus Language Contrasts

Activation Cluster	Extent (mm³)	Activation Peaks	x	y	z	Z Score
A. Language vs. environmental sound						
Left temporal	2,052	Left MTG (anterior)	−64	−26	0	3.10
		Left fusiform gyrus	−32	−44	−10	2.95
		Left MTG + STG (posterior)	−56	−38	6	2.78
Right temporal	2,376	Right ITG (posterior)	46	−60	−12	3.48
		Right ITG (posterior)	46	−54	−12	3.46
		Right STG (anterior)	56	−14	−6	3.32
		Right fusiform gyrus	40	−48	−18	3.06
		Right fusiform gyrus	38	−38	−24	2.97
		Right inferior occipital gyrus	50	−66	−16	2.68
		Right STG (polar)	58	6	−4	2.62
B. Environmental sound vs. language						
Right frontal	783	Right IFG (opercular)	44	10	32	3.63
		Right IFG (triangular)	50	28	20	3.32
Right temporal	2,295	Right STG (posterior)	56	−30	14	3.81
		Right STG (posterior)	58	−42	20	3.42
		Right STG (anterior)	64	−20	14	3.28
		Right STG (posterior)	64	−42	12	2.98

STG = superior temporal gyrus; ITG = inferior temporal gyrus.

with no significant difference over hemispheres. This was true when either an FDR-corrected threshold was used, as seen in Figure 4, or when a very liberal voxelwise $p < .05$ threshold was imposed (with FDR-corrected overlap numerically greater than with voxelwise thresholding). Exceptions to this rule were the opercular and particularly orbital sections of the right IFG, where there was little to no overlap when calculated on FDR-corrected volumes. However, at the voxelwise $p < .05$ threshold (where data from all subjects were included), overlap in these two areas was ∼40% and ∼60%, respectively. Only one sub-ROI, the right mid STS/STG/MTG, showed notably little overlap at even a very liberal threshold, with an average of only 11% overlap with FDR-corrected thresholds (including data from only four subjects) and 30% overlap when all subjects were included (voxelwise threshold of <.05).

It is important to note that there was a wide range of individual variability in overlap, with subjects showing almost complete overlap across ROIs (84%) to relatively little (28%). However, as demonstrated by Saad et al. (2003), the overlap in activation between two closely related tasks is positively correlated with power to detect activated voxels. Indeed, we found a positive rela-tionship between overall environmental sound/language overlap and overall volume of activation ($r = .625$, $p = .0299$, calculated on voxelwise $p < .05$ counts in order to include data from all subjects), showing that the greater volume of activation we observed in a subject, the more likely we were to find overlapping activation in the two domains. This suggests that nonoverlapping patches of activation may have stemmed from power considerations as well as from a true lack of shared neural resources.

Activation–Lesion Correlations

Our final set of analyses focused on the language-dominant LH alone. Here, we asked how well lesion maps derived from English-speaking patients with stroke (Saygin et al., 2003) could predict the extent and intensity of healthy subjects' fMRI activation in a given region.

In these lesion maps, each voxel's color shows how many patients with behavioral deficits have lesioned tissue in that voxel (e.g., see right side of Figure 5). The logic of a lesion map is as follows: If most or all impaired patients have lesions in a particular voxel, then
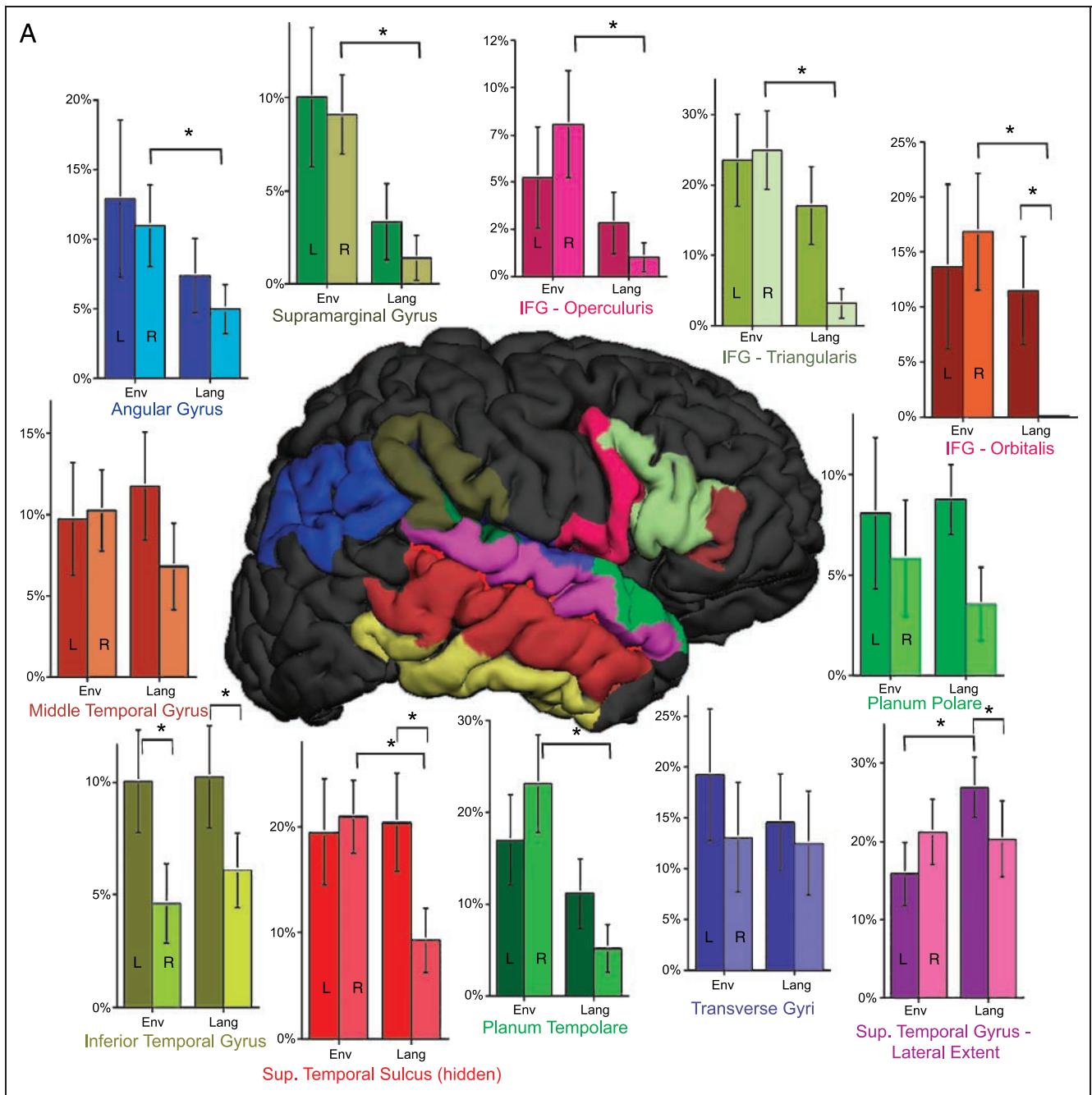
Figure 3. (A) Activation in each ROI for environmental sounds and language. For each graph, the *y*-axis shows the percent of voxels within each ROI that show suprathreshold positively signed activation (FDR-corrected *p* value; *p* < .05 for each ROI on a subject-by-subject basis). The color of each graph corresponds to the ROI projected onto a representative subject's pial surface. Statistically significant differences between conditions are indicated by the starred horizontal brackets above each graph. Error bars show ±1 *SEM*. L = left hemisphere; R = right hemisphere; IFG = inferior frontal gyrus; Sup. = superior; Env = environmental sounds; Lang = language. (B) Activation in the subparcellated superior and middle temporal gyri and sulci. Abbreviations and statistical differences as in (A).

the tissue in that voxel may be especially important for the task. Conversely, if none or few impaired patients have lesions in that voxel, then that tissue may be less important for performing the task (see Dronkers & Ogar, 2004; Saygin et al., 2003, for a more complete discussion).

Here we test whether the proportion of impaired subjects with lesioned tissue in a voxel predicts how active that voxel will be when healthy subjects perform the same task. For instance, we may predict higher levels of activation for voxels in regions where seven of eight environmental-sound (Env)-impaired patients have lesions (e.g., the posterior STG). Conversely, we may predict lower levels of activation for voxels in regions like the anterior middle frontal gyrus, where only one of eight Env-impaired patients have lesions. In other words
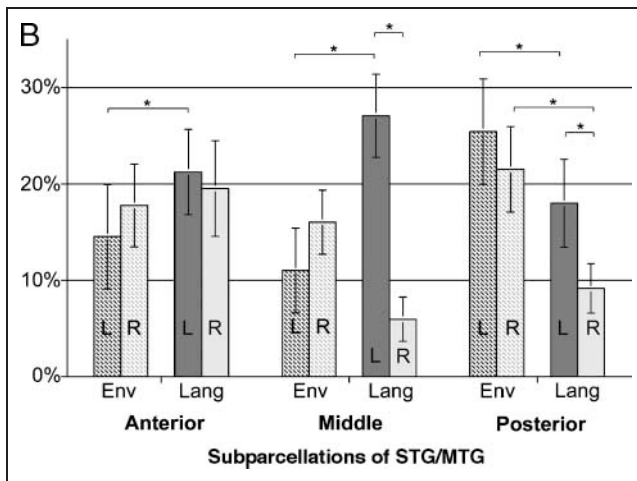
Figure 3. (continued)

we hypothesize that there should be a significant positive correlation between the number of Env-impaired patients with a lesion in a voxel or region, and the extent or intensity of MRI activation of Env-related fMRI activation in healthy subjects in that same voxel or region. (Figure 5 shows a direct comparison of lesion maps and activation overlaid on a single healthy subject from the fMRI experiment.)

To test this hypothesis, we calculated the average extent of fMRI activation (in percent of suprathreshold positively signed voxels) for each value or "level" of the lesion maps. For example, for all voxels where seven of eight Env-impaired patients had lesions, we calculated the percent of those voxels (averaged over subjects) that were active in healthy subjects performing the corresponding fMRI task. (This process was iterated for each other lesion map value, e.g., 6/8, 5/8, 4/8, etc.) We also

used the lesion maps to predict activation across domains, for example, using the Env lesion map to predict language (Lang) fMRI activation.

When we performed these analyses, we found predictive relationships between lesion maps and activation, both within and across domains. Using Spearman rank values (corrected for ties), we found that the number of Env-impaired patients with lesions per voxel (IPL/voxel) was predictive of the average percentage of positively signed suprathreshold voxels (% Active) in the environmental sounds condition ($\rho = .857$, $p = .0137$) and the language condition ($\rho = .786$, $p = .0362$) (see Figure 6A). Similarly, the IPL/voxel in the Lang lesion map also predicted the average percentage of positively signed suprathreshold voxels (% Active) in the language condition ($\rho = .771$, $p = .0724$) and the environmental sounds condition ($\rho = .943$, $p = .0048$), albeit with the former at marginal levels of significance (see Figure 6B). (In order to assure that results were not driven by just a few fMRI subjects, we repeated correlational analyses using a more lenient voxelwise threshold of $p < .01$ uncorrected for multiple comparisons; results did not change with the shift in activation threshold.)

We also examined the predictive relationship between lesion maps and fMRI activation when intensity rather than extent of activation was used as the dependent measure. Using the mean beta coefficient for each lesion-mask level as the measure of intensity, we found that the Env lesion map predicted intensity of both Env activation ($\rho = 1.00$, $p < .0001$) and Lang activation ($\rho = .75$, $p = .0522$) conditions (with the latter marginally significant). The Lang lesion map also predicted intensity of Env activation ($\rho = .943$, $p = .0048$), but did not significantly predict intensity of language activation ($\rho = $
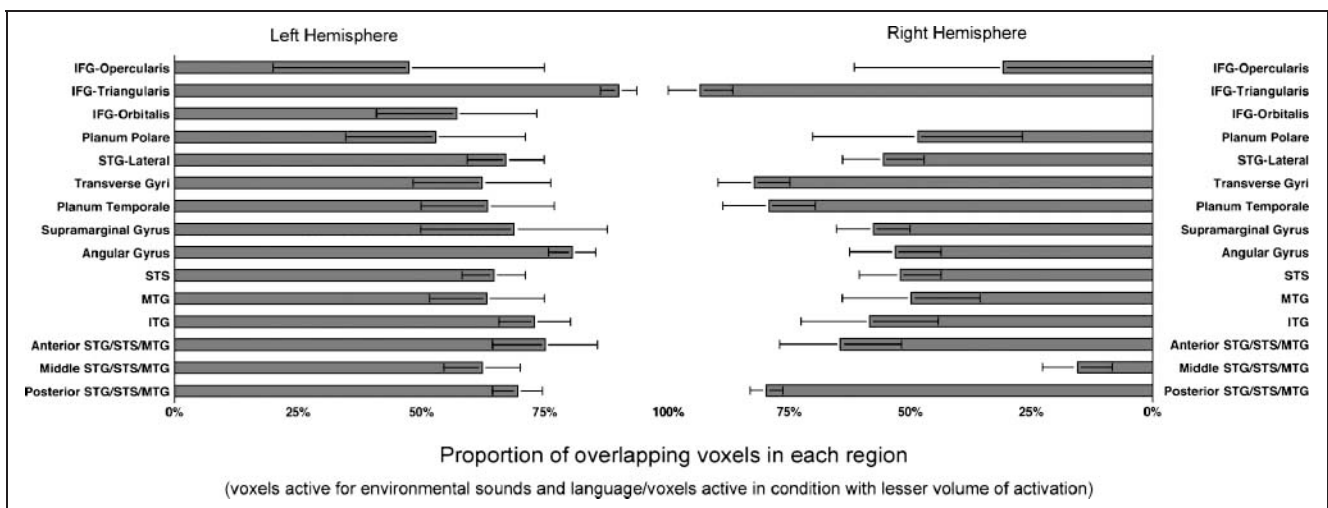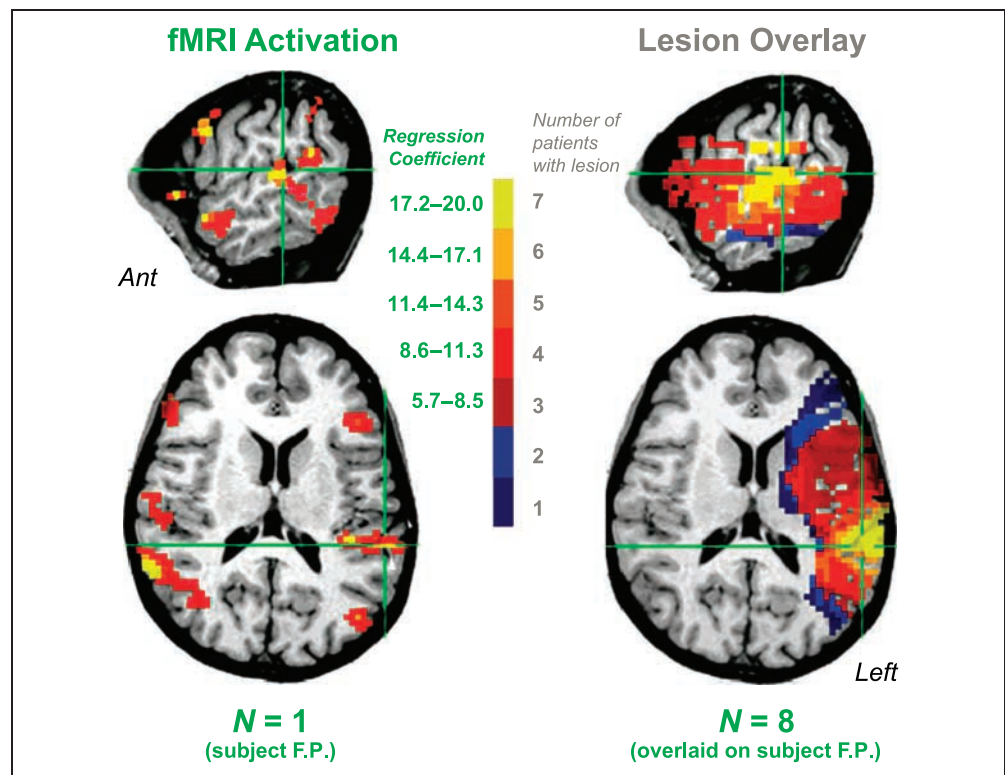

Figure 4. Overlap in activation between environmental sounds and language in each ROI. Each bar plots the overlap in activation between the two conditions, defined as the percentage of the smaller of the two activations falling within the larger one. Gray bars show overlap percentages at an FDR-corrected ROI-wise threshold of $p < .05$. Error bars indicate ± 1 *SEM*. IFG = inferior frontal gyrus; STG = superior temporal gyrus; MTG = middle temporal gyrus; ITG = inferior temporal gyrus; STS = superior temporal sulcus.

Figure 5. Left: A single subject's fMRI activation for environmental sounds. Activation was thresholded at voxelwise $p < .001$, with only positively signed voxels displayed; a voxel's color indicates the corresponding regression coefficient, as shown on the left side of the color scale. Right: Environmental sound lesion map, morphed to the same subject. Here, a voxel's color shows the number of patients with deficits who have lesions in that voxel (see right side of the color scale). For quantitative lesion–fMRI comparisons, see Figure 6A and B. Images are shown in radiological convention (right side of image = left side of brain). Ant = anterior.

fMRI Activation — Lesion Overlay

Regression Coefficient		Number of patients with lesion
17.2–20.0		7
14.4–17.1		6
11.4–14.3		5
8.6–11.3		4
5.7–8.5		3
		2
		1

Ant

Left

N = 1 (subject F.P.)   N = 8 (overlaid on subject F.P.)

.371, $p = .4062$). In general, the lower correlations between lesion maps and language activation—such as the lack of significant correlation between language lesion maps and intensity of language activation—were driven by the highly variable and overall lower levels of language activation in the left posterior STG, where the maximum number (6/10) of aphasic patients with language comprehension deficits had lesions.[2] When this region was eliminated from the analyses, language activation (both volume and intensity) was significantly or marginally correlated with both lesion maps.[3]

DISCUSSION

Results from these complementary analysis methods (whole-brain and ROI activation analyses, within-ROI overlap, and activation–lesion correlation) point toward a broadly shared and distributed cortical network of resources underlying both environmental sound and language processing, when these domains are compared using the same experimental task and semantically matched stimuli. This observed network was similar to that described in other recent studies of complex auditory processing, such as music (Koelsch et al., 2002) and audiovisual integration (Beauchamp, Lee, Argall, & Martin, 2004; Lewis et al., 2004). The different informational demands imposed by the two domains were reflected for the most part in broad quantitative shifts in activation across this network. Interestingly, the relative engagement of this distributed network was partic-

ularly similar over domains in the LH, both in terms of overall patterns of activation within particular ROIs, as well as in the correlation between lesion-symptom maps and extent and intensity of fMRI activation.

The cross-domain differences we did observe were generally in tune with those reported for comparisons between activation for speech and nonspeech sounds. For instance, in the LH, the whole-brain analyses showed a significant Lang > Env difference in the MTG and STS; the sub-ROI analyses confirmed that left anterior and particularly middle STG/STS/MTG were more active for language than environmental sounds. This pattern and location of differences was very similar to that reported by Scott, Blank, Rosen, and Wise (2000) for group differences between PET activation for speech (forward, vocoded, and reversed) and nonspeech (reversed vocoded) stimuli, although the overall center of activation appeared to be located slightly more posteriorly in the current study than in the Scott et al. study. These results were also evocative of a recent comparison of STS activation for human and nonhuman vocalizations: Relative to nonvocal sounds like music, Fecteau, Armony, Joanette, and Belin (2004) found that cat vocalizations drove more activation in a small portion of the left anterior STS, whereas human vocalizations (both speech and nonspeech) drove a broader and bilateral activation of this region, as well as more posterior aspects of the STS.

The disparity between language and environmental sound activation in the right posterior inferior temporal

Figure 6. Average percentage of suprathreshold voxels for the environmental sound (left) or language (right) fMRI task within the volume defined by all voxels with a given number of impaired patients with lesions (*x*-axis). ''No. of patients with lesions'' refers to the number of patients with environmental sound deficits (A) or language deficits (B) who showed a lesion in a particular voxel. ''% Active'' refers to the percent of the volume of the set of voxels (defined by IPL/voxel) that shows suprathreshold activation at an FDR-corrected value.

fMRI Correlations with Environmental Sounds Deficits

fmRI Correlations with Language Deficits
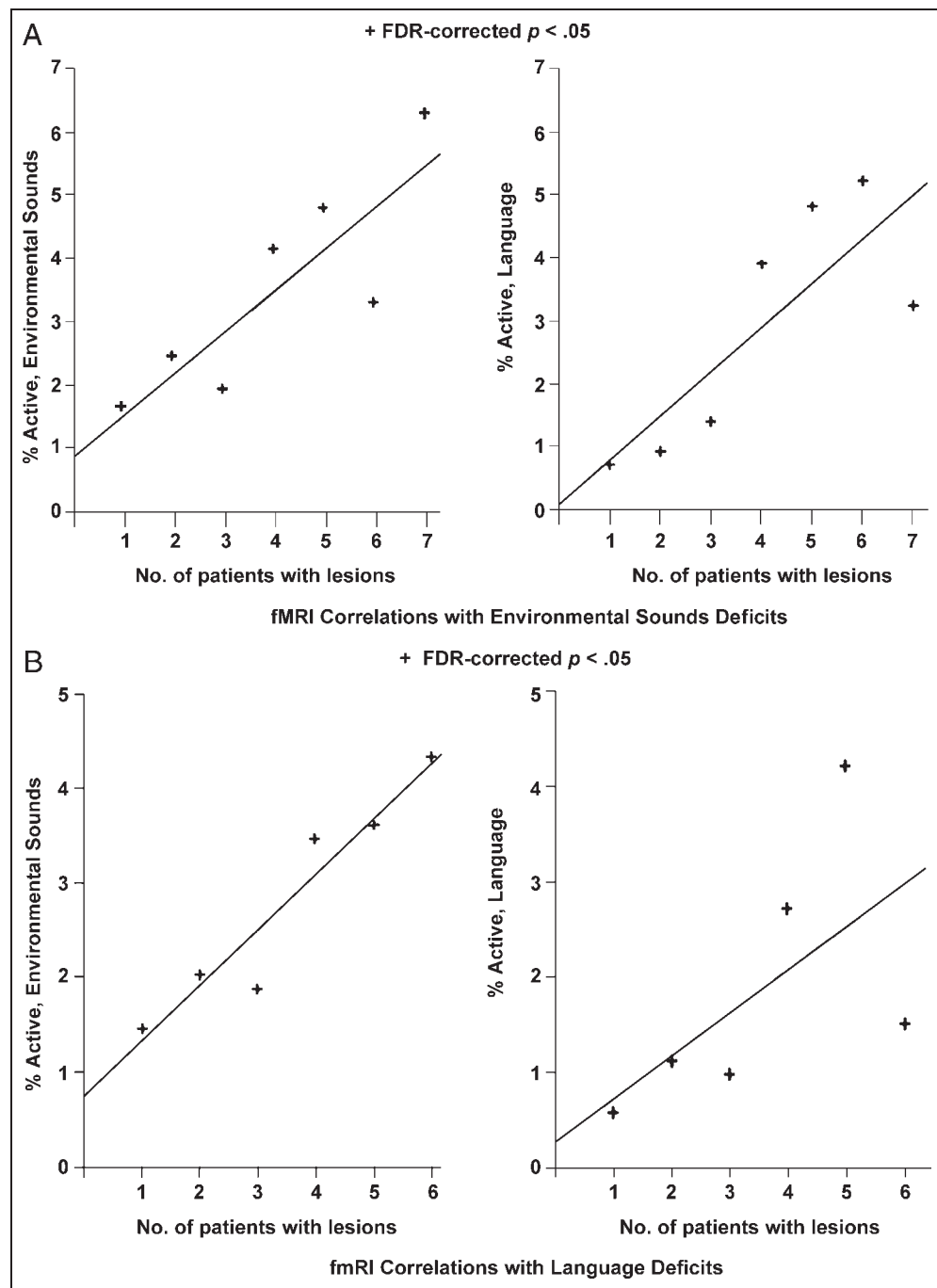
lobe is intriguing; if it had occurred on the *left*, it would have agreed with much literature about the ''basal temporal language area'' (Brodmann's area 37; see Büchel, Price, & Friston, 1998), which is often implicated in naming and reading tasks. The fact that language stimuli appeared to be evoking more activation in the *right* anterior medial fusiform area was somewhat unexpected, although there are hints of a similar effect in Humphries et al. (2001).

As revealed in the whole-brain group analyses, the greater activations for environmental sounds than for language in the right STG and planum temporale were

similar in terms of hemispheric lateralization to those reported by Belin, Zatorre, and Ahad (2002) and Belin et al. (2000). These studies suggested that speech or speechlike stimuli (with rapid temporal cues) preferentially modulated activation in the left temporal lobe, whereas nonspeech stimuli, particularly those relying on longer evolving frequency changes as in some of our environmental sound stimuli, preferentially modulated activation in the right temporal lobe (for a review, see Zatorre, Belin, & Penhume, 2002). It is possible that the lateralization differences between environmental sounds and language observed in some electrophysiological

studies (Plante et al., 2000; Van Petter & Rheinfelder, 1995) might also be driven by these acoustical differences (but cf. Cummings et al., 2006, who found no significant cross-domain differences in lateralization).

However, in looking at the results from the ROI analyses, it is important to highlight the fact that the two domains cannot easily be characterized as simply right or left lateralized—the data are more complicated, and more interesting. For instance, whereas language activation in the lateral STG was greater and more left lateralized than in the environmental sounds condition, the ITG had a very similar left-lateralized profile for both domains, and in the left posterior STG/STS/MTG, part of the classically defined Wernicke's area, environmental sounds evoked *more* activation than did these language stimuli. (More posteriorly, the supramarginal and angular gyri showed a marginally significant effect in the same direction. However, unlike Sprecht and Reul (2003), we did not find greater activation for environmental sounds in the left planum temporale and transverse gyri.)

Nor did we find that environmental sounds were processed dominantly in the RH, as some studies have suggested (Thierry et al., 2003): In our parcellation-based comparisons over hemisphere we found no significant right > left differences in extent of environmental sound activation (with only the planum temporale showing a marginally significant R > L advantage). What we *did* see is that environmental sounds tended to evoke a much more *bilateral* profile of activation. Because language was more left lateralized, there was generally more activation in right-hemisphere ROIs for environmental sounds than language, but in the right-hemisphere homologues of classic "language" areas, this was not always the case. Instead, we saw significantly more right-hemisphere activation for environmental sounds than for language in only about half the ROIs we measured (the inferior parietal lobe, the IFG, and the STS, and planum temporale). We did *not* see significant Env > Lang differences in more anterior and lateral regions of the temporal gyrus (e.g., planum polare, lateral STG, and transverse gyri), nor in the ITG—where in the whole-brain analyses we actually found *more* activation for language than for environmental sounds, as mentioned above. Although we do not want to overstate the case (by affirming the null hypothesis), it is certainly true that at least with our stimuli and task, there is no simple difference in asymmetry for the two domains.

The largest left–right disparities were in the IFG, where environmental sounds and language showed relatively equivalent activation in the LH ROIs, but language stimuli evoked relatively little activation in the right homologues, although environmental sounds strongly activated these regions. This is reminiscent of the results shown in Humphries et al. (2001). These laterality differences across domains may reflect the relative difference in "motor attention" devoted to the two domains, in that increases in motor attention tend

to go hand in hand with increases in left lateralization (Rushworth, Ellison, & Walsh, 2001). Whereas language is intimately connected with fine motor control of the vocal apparatus, many environmental sounds used in this study either cannot be produced by the human body, are the result of reflexive responses (like coughing), or are infrequently produced via an external instrument (like a piano or a violin). In this regard, our group (Pizzamiglio et al., 2005) has conducted a high-density EEG study to investigate whether perception of sounds referring to actions that can be performed by the perceiver might drive differential processing in the human brain. Here we used an audiovisual version of the repetition suppression paradigm to investigate the time course and locus of processing related to action-based sounds. Results showed that the left posterior superior temporal and premotor areas were selectively modulated by action-related sounds; in contrast, the temporal pole was bilaterally modulated by *non*-action-related sounds.

More striking in the current study were the overall similarities in activation over the LH—not only were all these "language-related" perisylvian ROIs as active overall for environmental sounds as for language (with the exception of the middle lateral STG/STS), but within-subject overlap in these regions was quite high, albeit with a good degree of individual variation. Finally, in this language-dominant hemisphere, we found that the degree to which a brain region was implicated in processing of environmental sounds in the lesions of a group of left-hemisphere-injured patients predicted the extent and intensity of functional activation for *both* environmental sounds and language in healthy young subjects. Further studies will be needed to explore the less straightforward relationship between lesion maps and fMRI activation within the language domain itself. It is also important to note that the lesion–fMRI comparison we have used here has its limits, particularly in terms of spatial resolution. We are currently developing more fine-grained approaches that rely on multivariate spatial statistics (Saygin, 2006); direct comparisons of the two methods should allow us to validate the present results.

In closing, our results suggest close links between the network of neural resources underlying the comprehension and/or recognition of language and complex, meaningful nonlinguistic stimuli. Observed differences were primarily of degree, rather than being absolute, particularly in the LH. This suggests that our language skills are subserved by brain areas that also contribute to nonlinguistic abilities that entail similar processing demands. Finally, lesion-deficit maps from both domains predicted the regional extent of fMRI activation in both domains, suggesting that the resources that are *necessary* for complex linguistic and nonlinguistic auditory processing are shown by functional imaging to be those that are also *involved* in processing this information in normally functioning brains.

Acknowledgments

Notes

1. Dick et al. (2002) compared RTs when subjects did or did not covertly name each sound in the same environmental sound–picture matching task, showing that covert naming significantly increased RTs. Environmental sound RTs in the present study were within 3 msec of those reported for the Dick et al. ''no-name'' condition, suggesting that subjects in the fMRI experiment were not covertly vocalizing during the environmental sound trials. Similar results are seen in the peak latencies of environmental sound and language event-related potentials (Cummings et al., 2006; Van Petten & Rheinfelder, 1995).

2. It is important to note that the behavioral correlations between environmental sounds and language processing were extremely high, so the difference between the two lesion masks is quite small.

3. Note that there is some risk of a false positive inherent in calculating multiple lesion map/fMRI correlations.

REFERENCES

Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *Journal of Neuroscience, 20,* 2583–2690.

Aziz-Zadeh, L., Iacoboni, M., Zaidel, E., Wilson, S., & Mazziotta, J. (2004). Left hemisphere motor facilitation in response to manual action sounds. *European Journal of Neuroscience, 19,* 2609–2612.

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 250–267.

Ballas, J. A., & Howard, J. H. (1987). Interpreting the language of environmental sounds. *Environment & Behavior, 19,* 91–114.

Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology, 40,* 293–310.

Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., et al. (2003). Voxel-based lesion-symptom mapping. *Nature Neuroscience, 6,* 448–450.

Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron, 41,* 809–823.

Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Brain Research, Cognitive Brain Research, 13,* 17–26.

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

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, 57,* 289–300.

Borovsky, A., Saygin, A. P., Cummings, A., Bates, E., Trauner, D., & Dick, F. (2006). Contrasting nonlinguistic and linguistic auditory processing in children with early focal lesions and language impairment. *Poster presented at the 12th Annual Conference on Architectures and Mechanisms for Language Processing, Nijmegen, the Netherlands.*

Bouvier, S. E., & Engel, S. A. (2006). Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral Cortex, 16,* 183–191.

Büchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway. *Nature, 394,* 274–277.

Casey, B. J., Tottenham, N., & Fossella, J. (2002). Clinical, imaging, lesion, and genetic approaches toward a model of cognitive control. *Developmental Psychobiology, 40,* 237–254.

Chiu, C., & Schacter, D. (1995). Auditory priming for nonverbal information: Implicit and explicit memory for environmental sounds. *Consciousness and Cognition, 4,* 440–458.

Clarke, S., Bellmann, A., De Ribaupierre, F., & Assal, G. (1996). Non-verbal auditory recognition in normal subjects and brain-damaged patients: Evidence for parallel processing. *Neuropsychologia, 34,* 587–603.

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, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25,* 257–271.

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

Cummings, A., Ceponiene, R., Koyama, A., Saygin, A. P., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds. *Brain Research, 1115,* 92–107.

Cummings, A., Dick, F., Saygin, A. P., Townsend, J., & Ceponiene, R. (2005). Electrophysiological responses to speech and environmental sounds in pre-adolescent and adolescent children. *Poster presented at the Society for Research on Child Development,* Atlanta, GA.

Cummings, A., Saygin, A. P., Bates, E., & Dick, F. (2005). The development of linguistic and non-linguistic auditory comprehension from 15–25 months: A preferential looking study. *Manuscript submitted for publication.*

Cycowicz, Y. M., & Friedman, D. (1998). Effect of sound familiarity on the event-related potentials elicited by novel environmental sounds. *Brain and Cognition, 36,* 30–51.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage, 9,* 179–194.

DeArmond, S. J., Fusco, M. M., & Dewey, M. M. (1976). *Structure of the human brain: A photographic atlas.* New York: Oxford University Press.

Dick, F., Bussiere, J., & Saygin, A. P. (2002). The effects of linguistic mediation on the identification of environmental sounds. *Center for Research in Language Newsletter, 14,* 3–9.

Dronkers, N., & Ogar, J. (2004). Brain areas involved in speech production. *Brain, 127,* 1461–1462.

Duvernoy, H. M. (1999). *The human brain: Surface, three-dimensional sectional anatomy with MRI, and blood supply.* Wien: Springer.

Fecteau, S., Armony, J. L., Joanette, Y., & Belin, P. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study. *Neuroimage, 23,* 840–848.

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

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

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

Friedman, D., Cycowicz, Y. M., & Dziobek, I. (2003). Cross-form conceptual relations between sounds and words: Effects on the novelty P3. *Cognitive Brain Research, 18,* 58–64.

Frey, R. T., Woods, D. L., Knight, R. T., Scabini, D., & Clayworth, C. (1987). Defining functional areas with averaged CT scans. *Society of Neuroscience Abstracts, 13,* 1266.

Friston, K., Ashburner, J., Poline, J., Frith, C., Heather, J., & Frackowiak, R. (1995). Spatial registration and normalization of images. *Human Brain Mapping, 2,* 165–189.

Friston, K. J., Holmes, A., Poline, J. B., Price, C. J., & Frith, C. D. (1996). Detecting activations in PET and fMRI: Levels of inference and power. *Neuroimage, 4,* 223–235.

Friston, K. J., Holmes, A., & Worsley, K. J. (1999). How many subjects constitute a study? *Neuroimage, 10,* 1–5.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage, 15,* 870–878.

Gentilucci, M. (2003). Object motor representation and language. *Experimental Brain Research, 153,* 260–265.

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.

Gygi, B., Kidd, G. R., & Watson, C. S. (2004). Spectral–temporal factors in the identification of environmental sounds. *Journal of the Acoustical Society of America, 115,* 1252–1265.

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., et al. (1999). ''Sparse'' temporal sampling in auditory fMRI. *Human Brain Mapping, 7,* 213–223.

Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41,* 301–307.

Hillis, A. E., Work, M., Barker, P. B., Jacobs, M. A., Breese, E. L., & Maurer, K. (2004). Re-examining the brain regions crucial for orchestrating speech articulation. *Brain, 127,* 1479–1487.

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.

Jäncke, L., Buchanan, T., Lutz, K., Specht, K., Mirzazade, S., & Shah, N. J. (1999). The time course of the BOLD response in the human auditory cortex to acoustic stimuli of different duration. *Brain Research, Cognitive Brain Research, 16,* 117–124.

Knight, R. T., Scabini, D., Woods, D. L., & Clayworth, C. (1988). The effects of lesions of superior temporal gyrus and inferior parietal lobe on temporal and vertex components of the human AEP. *Electroencephalography and Clinical Neurophysiology, 70,* 499–509.

Koelsch, S., Gunter, T. C., von Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical ''language-network'' serves the processing of music. *Neuroimage, 17,* 956–966.

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences, 9,* 578–584.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25,* 259–284.

Lazar, N. A., Luna, B., Sweeney, J. A., & Eddy, W. F. (2002). Combining brains: A survey of methods for statistical pooling of information. *Neuroimage, 16,* 538–550.

Lewis, J. W., Brefczynski, J. A., Phinney, R. E., Janik, J. J., & DeYoe, E. (2005). Distinct cortical pathways for processing tool versus animal sounds. *Journal of Neuroscience, 25,* 5148–5158.

Lewis, J. W., Wightman, F. L., Brefczynski, J. A., Phinney, R. E., Binder, J. R., & DeYoe, E. A. (2004). Human brain regions involved in recognizing environmental sounds. *Cerebral Cortex, 14,* 1008–1021.

MacWhinney, B. (1999). *The emergence of language.* Mahway, NJ: Erlbaum.

Moses, P., & Stiles, J. (2002). The lesion methodology: Contrasting views from adult and child studies. *Developmental Psychobiology, 40,* 266–277.

Ozcan, M., Baumgartner, U., Vucurevic, G., Stoeter, P., & Treede, R. D. (2005). Spatial resolution of fMRI in the human parasylvian cortex: Comparison of somatosensory and auditory activation. *Neuroimage, 25,* 877–887.

Pizzamiglio, L., Aprile, T., Spitoni, G., Pitzalis, S., Bates, E., D'Amico, S., et al. (2005). Semantic processing of action/non-action related sounds. *Neuroimage, 24,* 852–861.

Plante, E., van Petten, C., & Senkfor, A. J. (2000). Electrophysiological dissociation between verbal and nonverbal semantic processing in learning disabled adults. *Neuropsychologia, 38,* 1669–1684.

Rushworth, M. F., Ellison, A., & Walsh, V. (2001). Complementary localization and lateralization of orienting and motor attention. *Nature Neuroscience, 4,* 656–661.

Saad, Z. S., Ropella, K. M., DeYoe, E. A., & Bandettini, P. A. (2003). The spatial extent of the BOLD response. *Neuroimage, 19,* 132–144.

Saccuman, C., Dick, F., Bates, E., Mueller, R. A., Bussiere, J., Krupa-Kwiatkowski, M., et al. (2002). Lexical access and sentence processing: A developmental fMRI study. *Poster presented at the Meeting of the Cognitive Neuroscience Society, San Francisco, CA.*

Saygin, A. P. (2006). Combining results from functional MRI and voxel-based lesion mapping to study biological motion processing. *Presented at the European Conference on Visual Perception, St. Petersburg, Russia.*

Saygin, A. P., Dick, F., & Bates, E. (2005). An online task for contrasting auditory processing in the verbal and nonverbal domains and norms for college-age and elderly subjects. *Behavior Research Methods, 37,* 99–110.

Saygin, A. P., Dick, F., Wilson, S. W., Dronkers, N. F., & Bates, E. (2003). Shared 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.

Specht, K., & Reul, J. (2003). Functional segregation of the temporal lobes into highly differentiated subsystems for auditory perception: An auditory rapid event-related fMRI-task. *Neuroimage, 20,* 1944–1954.

Stuart, G. P., & Jones, D. M. (1995). Priming the identification of environmental sounds. *Quarterly Journal of Experimental Psychology, 48,* 741–761.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience, 17,* 273–281.

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

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15,* 273–289.

Van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia, 33,* 485–508.

Varney, N. R. (1980). Sound recognition in relation to aural language comprehension in aphasic patients. *Journal of Neurology, Neurosurgery, & Psychiatry, 43,* 71–75.

Worsley, K. J., Marrett, S., Neelin, P., Friston, K., & Evans, A. (1995). *A unified statistical approach for determining significant signals in images of cerebral activation.* San Diego, CA: Academic Press.

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