

Distinct Cortical Pathways for Processing Tool versus Animal Sounds

James W. Lewis,^{1,2} Julie A. Brefczynski,⁴ Raymond E. Phinney,^{3,6} John J. Janik,^{3,5} and Edgar A. DeYoe³

¹Department of Physiology and Pharmacology and ²Center for Advanced Imaging, West Virginia University, Morgantown, West Virginia 26506, Departments of ³Radiology and ⁴Cell Biology, Neurobiology, and Anatomy, Medical College of Wisconsin, Milwaukee, Wisconsin 53226, ⁵Department of Biomedical Engineering, Marquette University, Milwaukee, Wisconsin 53201, and ⁶Department of Psychology, Wheaton College, Wheaton, Illinois 60187

Human listeners can effortlessly categorize a wide range of environmental sounds. Whereas categorizing visual object classes (e.g., faces, tools, houses, etc.) preferentially activates different regions of visually sensitive cortex, it is not known whether the auditory system exhibits a similar organization for different types or categories of complex sounds outside of human speech. Using functional magnetic resonance imaging, we show that hearing and correctly or incorrectly categorizing animal vocalizations (as opposed to hand-manipulated tool sounds) preferentially activated middle portions of the left and right superior temporal gyri (mSTG). On average, the vocalization sounds had much greater harmonic and phase-coupling content (acoustically similar to human speech sounds), which may represent some of the signal attributes that preferentially activate the mSTG regions. In contrast, correctly categorized tool sounds (and even animal sounds that were miscategorized as being tool-related sounds) preferentially activated a widespread, predominantly left hemisphere cortical “mirror network.” This network directly overlapped substantial portions of motor-related cortices that were independently activated when participants pantomimed tool manipulations with their right (dominant) hand. These data suggest that the recognition processing for some sounds involves a causal reasoning mechanism (a high-level auditory “how” pathway), automatically evoked when attending to hand-manipulated tool sounds, that effectively associates the dynamic motor actions likely to have produced the sound(s).

Key words: auditory cortex; fMRI; association cortex; multisensory; mirror system; motor processing

Introduction

At an early age, we begin to learn to categorize different objects and their associated sounds into different semantic or conceptual categories. The idea that constructive processes operate to create conceptual categories in our minds has been around since the days of the early Greek (Medin and Coley, 1998) and Buddhist (Lusthaus, 2002) philosophers. In general, our ability to recognize different types of environmental (nonverbal) sounds is thought to be accomplished, in part, by abstracting the physical properties of a sound and matching them to the known characteristics of a given sound category (Komatsu, 1992; Medin and Coley, 1998). For instance, animal vocalizations often contain frequencies and harmonics that covary in time, which may serve as signal attributes or compound cues for recognition (Nelken et al., 1999; Reide et al., 2001; Ehret and Riecke, 2002). Although less well studied, some sounds produced by tools may share common

acoustical features, such as a metallic twang or high-pitched ringing sound, which could conceivably serve as “low-level” signal attributes that aid in our ability to recognize them as tools. Of course, the degree to which a person can specifically identify a given sound (such as recognizing the distinctive clicking sounds of a $\frac{3}{8}$ vs $\frac{1}{2}$ inch ratchet wrench) depends primarily on the listener’s experience or expertise. Thus, cognitive and other “high-level” associations are also likely to play a key role in the process of sound recognition (Kéri, 2003; Thompson-Schill, 2003).

More recently, human lesion and neuroimaging studies have indicated that the retrieval of words and other aspects of conceptual knowledge pertaining to distinct object categories are represented, at least in part, along separate brain regions (De Renzi and Lucchelli, 1988; Gainotti et al., 1995; Damasio et al., 1996; Martin et al., 1996; Chao et al., 1999; Moore and Price, 1999; Chao and Martin, 2000; Martin, 2001; Grossman et al., 2002; Ilmberger et al., 2002). Using visual or verbal input, the conceptual categories most consistently revealing segregated cortical pathways or networks include “man-made items” versus “living things.” A few studies have examined brain responses to different categories of sound, suggesting, for instance, that human speech sounds are processed differently from other sound categories, such as musical instruments, tones, environmental sounds, and animal cries (Belin et al., 2000; Binder et al., 2000; Levy et al., 2001; Fecteau et al., 2004). However, to our knowledge, no studies have systematically contrasted a wide array of sounds representing the conceptual categories of man-made items (tools in use) versus living

Received Aug. 9, 2004; revised April 17, 2005; accepted April 18, 2005.

This work was supported by National Institutes of Health (NIH) Grants R03 DC04642 (J.W.L.) and EY10244 and MH51358 (E.A.D.) and General Clinical Research Center/NIH Grant M01-RR00058 (Medical College of Wisconsin). We thank B. Doug Ward for assistance with paradigm design and statistical analyses, Nathan Walker for assistance with error analyses and figures, Dr. Harold Burton for comments on the paradigm, Dr. Kristina Ropella for acoustical analysis suggestions, and Dr. Aina Puce for constructive comments on previous versions of this manuscript. We thank Dr. Robert Cox for continual development of AFNI and Dr. David Van Essen, Donna Hanlon, and John Harwell for continual development of cortical data presentation with CARET.

Correspondence should be addressed to Dr. James W. Lewis, Department of Physiology and Pharmacology, P.O. Box 9229, West Virginia University, Morgantown, WV 26506. E-mail: jlewis@hsc.wvu.edu.

DOI:10.1523/JNEUROSCI.0419-05.2005

Copyright © 2005 Society for Neuroscience 0270-6474/05/255148-11\$15.00/0

things (animal sounds). Thus, using functional magnetic resonance imaging (fMRI), we sought to determine whether hearing and categorizing “tool” versus “animal” sounds would selectively activate different cortical pathways, thereby revealing a high-level architectural and functional organization of the human auditory system for sound recognition (Lewis and DeYoe, 2003).

Materials and Methods

Participants and main paradigm. We tested 20 right-handed adult participants (aged 21–52 years; 10 women) with no history of neurologic, psychiatric, or audiologic symptoms. Informed consent was obtained following guidelines approved by the Medical College of Wisconsin Human Research Review Committee. Participants (with eyes closed) were presented with 306 stimulus trials (108 animal sounds, 108 tool sounds, and 90 silent trials, over six separate fMRI scans) in a random order. They were informed that they would hear a large selection of animal and tool sounds, together with silent periods. Their task was to listen to the sounds and respond silently in their heads as to whether the likely sound-source was a tool or an animal. Participants were explicitly informed not to make any overt motor or vocal responses while being scanned.

Tool- and animal-related sound samples (44.1 kHz, 16-bit, stereo) were compiled from compact disk collections of professionally recorded sounds (Sound Ideas, Richmond Hill, Ontario, Canada). Sounds were typically recorded using stereo microphones, which contained two directional monaural microphones with 90° to 120° separation. The tool-related sounds were all from items typically manipulated by one hand (excluding motorized power tools; see Appendix, Table A), and the animal sounds were predominantly vocalizations (Appendix, Table B). The subsets of tool versus animal sounds were carefully matched overall (Cool Edit Pro; Syntrillium Software, Phoenix, AZ) for average root mean squared (RMS) power (loudness) and duration (1.9–2.1 s), and both stimulus sets contained a wide range of frequencies, tempos, and cadences, including sounds that were nearly continuous and those containing 1–14 distinct epochs (e.g., a dog producing four asynchronous barks). An envelope of 20 ms rise and fall times was applied to all sounds to minimize clicks at sound onset and offset. Human speech sounds (conversations, calls, exclamations, etc.) were also compiled as described above (Sound Ideas) but were only used for the purpose of comparing acoustic signal properties, as addressed below.

Imaging and data analysis. Scanning was conducted at 1.5 T on a General Electric (GE Medical Systems, Milwaukee, WI) Signa scanner, equipped with a commercial head coil (Medical Advances, Milwaukee, WI) suited for whole-head, echo planar imaging of blood-oxygenation level-dependent (BOLD) signals (Bandettini et al., 1993). We used a “silent” clustered-acquisition fMRI design that allowed sound stimuli (typically 70–85 dB, L-weighted through ear plugs) to be presented during scanner silence, over electrostatic headphones (Koss Inc, Milwaukee, WI). The scanning cycle has been described previously (Lewis et al., 2004). Briefly, every 10 s, a sound (or silent event) was presented, and, 7.5 s after stimulus onset, there followed the collection of BOLD signals from axial brain slices (echo time, 40 ms; repetition time, 1.8 s; 52 gradient-recalled image volumes). Image volumes included 16–18 axial slices of 6 mm thickness, with in-plane voxel dimensions of 3.75 × 3.75 mm. T1-weighted anatomical MR images were collected using a spoiled Grass Instruments (Quincy, MA) pulse sequence (1.1 mm slices, with 0.9375 × 0.9375 mm in-plane resolution). Immediately after the scanning session, participants heard all of the sounds again and indicated whether they judged each sound (when initially heard in the scanner) as either (1) an animal or (2) a tool, or if they were (3) uncertain as to the category of the sound source.

Data were viewed and analyzed using AFNI and related software plugins (<http://afni.nimh.nih.gov/>) (Cox, 1996). For each participant, the six scans were concatenated into one time series. Brain volume images were motion corrected for global head translations and rotations by re-registering them to the 20th brain volume of the last scan (closest to the time of anatomical image acquisition). Multiple linear regression analyses compared the BOLD brain responses with tool versus animal sounds, both relative to silence. Only those sounds correctly categorized were retained for the main analysis (see Fig. 1): BOLD responses to sounds that

were later found to have been incorrectly categorized or deemed as uncertain by each participant could be censored (excluded) from the deconvolution model for that individual (see parentheses in Appendix, Tables A, B). Additionally, after data collection, 14 animal sounds that were not clearly vocalizations and 14 tool sounds that were not strongly associated with use by the dominant hand were subsequently censored from all of the data analyses.

Individual anatomical and functional brain maps were transformed into the standardized Talairach coordinate space (Talairach and Tournoux, 1988). Functional data (multiple regression coefficients) were spatially low-pass filtered (5 mm box filter) and then merged by combining coefficient values for each interpolated voxel across all participants. The combination of individual voxel probability threshold (*t* test; $p < 0.005$) and the cluster size threshold (8 voxel minimum) yielded the equivalent of a whole-brain corrected significance level of $\alpha < 0.05$. A split-half correlation test yielded a voxel-by-voxel correlation between two subgroups (matched for gender and age) of 0.79, resulting in a Spearman-Brown estimated reliability coefficient of $\rho_{XY} = 0.88$ for the entire sample of 20 participants (Binder et al., 1997; Lewis et al., 2004). This indicates the level of correlation that would be expected between the activation pattern of our sample of 20 participants and activation patterns from other random samples of 20 participants matched for gender, age, and handedness. Public domain software packages SureFit and Caret (<http://brainmap.wustl.edu>) were used to project data onto the Colin Atlas brain (AFNI-*tlrc*) and to display the data (Van Essen et al., 2001; Van Essen, 2003). Portions of these data can be viewed at <http://sumsdb.wustl.edu:8081/sums/directory.do?dirid=707082>, which contains a database of surface-related data from other brain mapping studies.

Virtual tool manipulation task. Twelve of the 20 participants (six female) also performed a “virtual” tool manipulation task. They alternated (five cycles, 20 s blocks, eyes closed) between resting their hand versus making hand and arm movements (distal to the elbow) as if manipulating a variety of different tools, although no tools were actually being held. This included three to five different virtual tools from the list in Appendix (Table A), typically including hammering, sawing, scissoring, sanding, and ratcheting. However, other pantomimes were included for some participants after the first scan if deemed necessary to minimize incidental head movements, which were closely monitored in real time. Blocks were cued by different tone pips, and no other sound stimuli were presented. Using methods described previously (Lewis et al., 2000), three to four scan repetitions were averaged, and the time series data were then cross-correlated with “ideal” sinusoid reference waveforms. Response magnitudes were calculated as the amplitude of the best-fit reference waveform. Individual maps were transformed into Talairach space coordinates, spatially low-pass filtered (5 mm box filter), and then merged across participants (*t* test; $p < 0.005$; whole-brain corrected significance level of $\alpha < 0.05$).

Spectral and temporal analyses of sounds. Spectrographs, amplitude plots, individual power plots, and harmonics-to-noise-ratio (HNR) analyses of the tool, animal, and speech sounds were generated using freely available phonetic software (Praat; <http://www.fon.hum.uva.nl/praat/>). The HNR algorithm (below) determined the degree of periodicity within a sound signal, $x(t)$, based on finding a maximum autocorrelation, $r'_x(\tau_{\max})$, of the signal at a time lag (τ) greater than zero (Boersma, 1993):

$$\text{HNR (in dB)} = 10 \cdot \log_{10} \frac{r'_x(\tau_{\max})}{1 - r'_x(\tau_{\max})}$$

This measure quantified the acoustic energy of the harmonics that were present within a sound over time, $r'_x(\tau_{\max})$, relative to that of the remaining “noise,” $1 - r'_x(\tau_{\max})$, which represents nonharmonic, irregular, or chaotic acoustic energy. As extreme examples, a 2 s sample of white noise yielded an HNR value of -7.6 (using standard parameters: 10 ms time step; 75 Hz minimum pitch; 1 period per window), whereas a sample consisting of 2 and 4 kHz sine-wave tones produced an HNR value of 65.4.

For the averaged power spectra plots, we analyzed power (in Hertz) of the frequency spectrum (0–22 kHz, 1 Hz increments), which was estimated using the Thomson multi-taper method with a sampling frequency of 44.1 kHz (Matlab; MathWorks, Natick, MA). The power spectrum of each

sound was normalized to the total power of the respective sound. The normalized power spectra were summed (94 per sound category), and each frequency was presented as a proportion of the total power.

The bicoherence analysis (a form of bispectral analysis; Matlab) quantified the degree of phase coupling between all frequency pairs (f_1 and f_2 , at 1 Hz increments), examining linear and nonlinear second-order interactions of acoustic components (Sigl and Chamoun, 1994). In the bispectral equation below, $X_i(f_1)$ represents the power at f_1 (its Fourier transform, which contains a real and imaginary component), and X_i^* represents the complex conjugate of the power:

$$B(f_1, f_2) = \left| \sum_{i=1}^L X_i(f_1) X_i(f_2) X_i^*(f_1 + f_2) \right|.$$

The bispectrum calculation (the resulting real triple product) was normalized to create a bicoherence measure, thereby producing a measure of phase coupling independent of amplitude. Larger values reflected related, as opposed to random, associations between frequency pairs over time. The bicoherence analysis yielded a two-dimensional (2-D) rendering with a fourfold symmetry about cardinal axes (data not shown), which were averaged across all 94 sounds within a category. To aid in the visualization of differences in the 2-D plots, the real component (the first octant triangle) was collapsed across one frequency dimension (f_2) to create a one-dimensional bicoherence “profile” for each category of sound.

Results

Tool versus animal sound paradigm

We scanned 20 right-handed participants while they listened to a random presentation of tool sounds (Appendix, Table A; hand-manipulated tools in use), animal sounds (Appendix, Table B; predominantly vocalizations), and silent events. Participants were instructed to categorize each stimulus as either a tool or animal sound silently in their head. Participants were explicitly instructed not to produce any motor responses (i.e., button presses, subvocally naming sounds, etc.) during the scan session so as to avoid confounding any activation attributable to sound processing with that attributable to motor output responses. We examined focal brain activation (changes in fMRI signals) to correctly and incorrectly categorized sound stimuli based on response data collected immediately after the scanning session (on average, 92% correct, 4% miscategorized, and 4% uncertain).

The group-averaged pattern of activation attributable to hearing and correctly categorizing both tool and animal sounds relative to silence included a widespread pattern of activation in both hemispheres (Fig. 1*a*, yellow to orange). As expected, the strongest activation (yellow; $p < 0.00005$ threshold) included primary auditory cortex (overlapping Heschl's gyrus) plus other auditory regions along the planum temporale and superior temporal plane (collectively termed “PAC+”), located within the lateral sulci (LaS) (Wessinger et al., 2001). Moderate activation (orange–yellow, $p < 0.0001$; and orange, $p < 0.001$; $\alpha < 0.05$) was present along portions of the superior temporal gyri (STG) and superior temporal sulci (STS), the left inferior frontal region (precentral sulcus), the superior frontal gyri on the medial wall (not visible in lateral views), and in subcortical structures, including bilateral thalamic and caudate nuclei (data not shown). The activation pattern attributable to hearing and correctly categorizing tools (Fig. 1*b*, red) relative to silence and animals (Fig. 1*c*, blue) relative to silence revealed a roughly similar overall pattern of activation, although the tools activated a more expansive network.

In contrast, some regions of cortex (Fig. 1*a–c*, light green) showed either a depression below baseline in response to the sound stimuli or a relatively greater activation during the silent periods, which may in part be attributable to task-unrelated ef-

fects, such as semantic monitoring, day dreaming, or other thought processes during silent trials (Binder et al., 1999; Calvert and Lewis, 2004). These regions included much of the dorsal occipital cortex and portions of dorsal frontal cortex in both hemispheres.

To reveal brain regions preferentially involved in processing tool-related sounds relative to animal sounds, and vice versa, we effectively subtracted (via multiple linear regression analysis on an individual basis) the activation pattern for the animal sounds versus silence (Fig. 1*c*) from that for the tool sounds versus silence (Fig. 1*b*), both at zero threshold. The resulting pattern of activation is illustrated on lateral views (Fig. 1*d*) and the corresponding flat-map models (Fig. 1*e*) of the Colin Atlas brain and on select axial slices (Fig. 2) from one of the participants. In these maps, red depicts cortex that was preferentially activated by the tool sounds and blue depicts regions preferentially activated by the animal sounds. No significant differential activity to tool versus animal sounds was observed in subcortical brain regions, and no significant differences were observed across gender ($\alpha < 0.05$ in a two-sample *t* test for means).

The main finding from the tool versus animal sound comparison (Figs. 1*d,e*, 2) was that several cortical regions were differentially activated by the two different categories of sound. Animal sounds evoked significantly stronger activity (blue) in both the left and right hemisphere (bilaterally) along middle portions of the STG (mSTG). Importantly, this bias was present in all 20 individuals. In contrast, tool sounds evoked activity (red) mostly in the left hemisphere. This included nine major cortical foci (Fig. 1): (1) the middle portion of the left inferior frontal sulcus (mIFS); (2) the left ventral premotor cortex (VPMC); (3) the left inferior postcentral plus temporo-parietal junction and parietal opercular cortex (collectively referred to as IPoCeC); (4, 5) posterior portions of the left and right lateral sulci (pLaS); (6) the left anterior intraparietal regions, overlapping previously defined area “AIP” (Binkofski et al., 1999; Buccino et al., 2001; Grèzes et al., 2003); (7) portions of the left posterior parietal cortex; and (8, 9) the left and right posterior middle temporal gyri (pMTG) (including posterior portions of the STS).

Note that, in our previous study using 15 of the 20 same participants, the bilateral mSTG foci (blue) for animal sounds were found to be relatively insensitive to the “perceived recognition” of natural sounds per se: these foci were comparably activated by both recognizable environmental sounds (a more diverse range of sound categories) as well as by the corresponding backward-played versions of the same sounds, which were judged as unrecognizable (Lewis et al., 2004). In contrast, the bilateral pMTG foci (red) plus portions of the left IPoCeC, left VPMC and mIFS foci preferential for tool sounds directly overlapped cortex implicated in environmental sound “recognition” in our previous study.

Region of interest analysis and miscategorized sounds

To further assess the preferential activations attributable to tool versus animal sounds and their representation as distinct “conceptual” categories, we charted the averaged BOLD response magnitudes (Fig. 2, charts) within 11 cortical regions-of-interest (ROIs) from Figure 1, *d* and *e*. The first column depicts responses to correctly categorized tool sounds (“T,” red) relative to silence, and the second column depicts the responses to correctly categorized animal sounds (“A,” dark blue) relative to silence. Four of the ROIs showed increased responses to tool sounds but showed decreased responses to animal sounds, including the left and right pMTG, left posterior parietal, and left AIP cortex. These response characteristics may have been indicative of processing that was

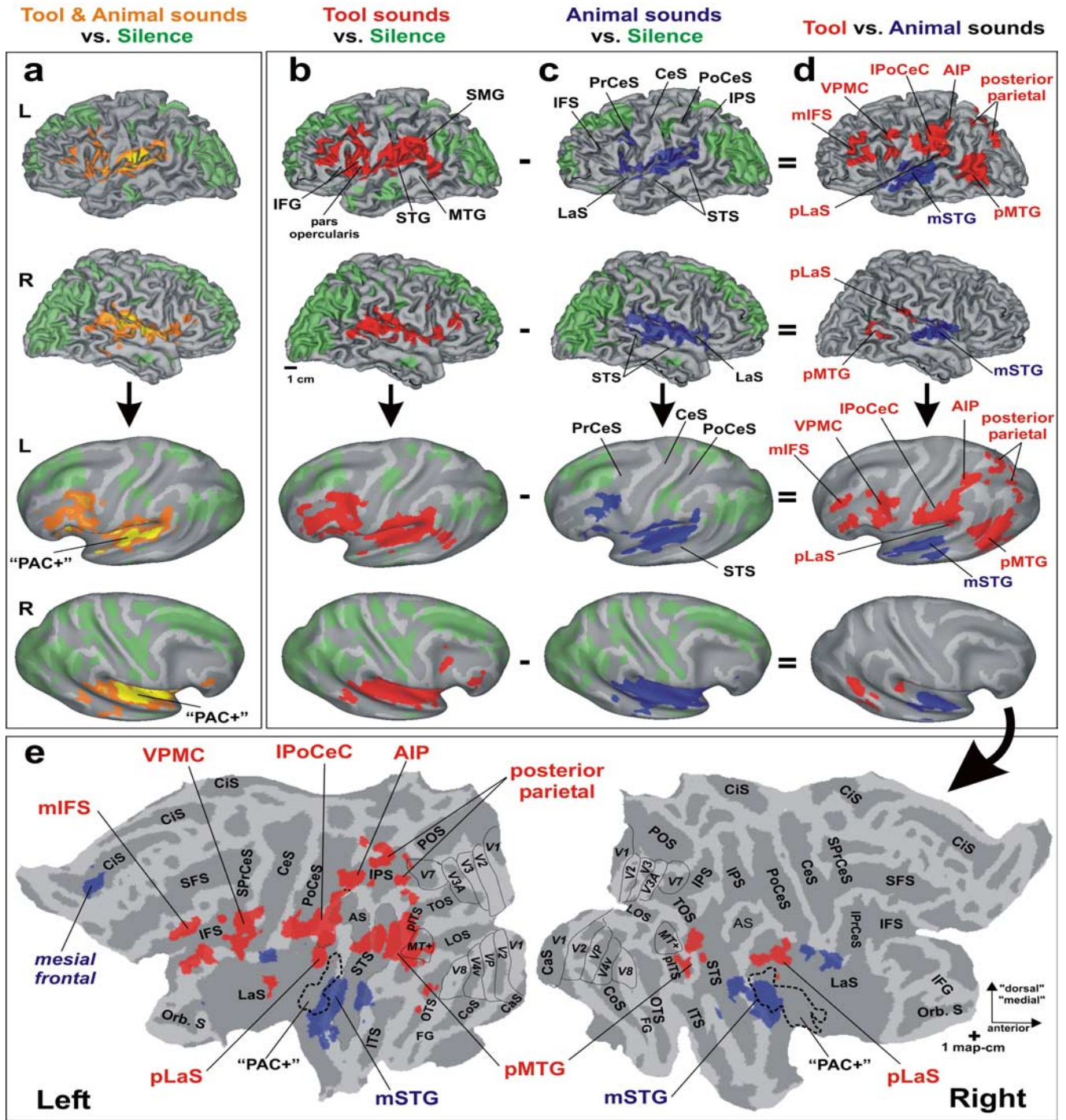


Figure 1. BOLD responses to correctly categorized tool and animal sounds ($n = 20$; all panels corrected to $\alpha < 0.05$). Cortical regions activated by both tool and animal sounds (yellow to orange) relative to silence (**a**), tool-related sounds (red) relative to silence (**b**), and animal vocalization sounds (blue) relative to silence (**c**). Light green depicts cortex showing decreased BOLD signal relative to silence. **d**, Data from **c** effectively subtracted from **b**, revealing regions preferentially activated by tool sounds (red) or by animal sounds (blue). Bottom panels depict highly smoothed renderings of the Colin Atlas brain. **e**, Flat maps showing data from **d**. Black dashed outlines depict the approximate location of PAC+ from **a** (yellow; $p < 0.00005$). Identified visual areas (solid black outlines; V1, V2, V3, MT+, etc.) are from the Colin Atlas database. For details, see Results and Table 1. AS, Angular sulcus; CaS, calcarine sulcus; CeS, central sulcus; CIS, cingulate sulcus; CoS, collateral sulcus; FG, fusiform gyrus; IFG, inferior frontal gyrus; IFS, inferior frontal sulcus; IPrCeS, inferior precentral sulcus; IPS, intraparietal sulcus; ITS, inferotemporal sulcus; LaS, lateral sulcus; LOS, lateral occipital sulcus; MTG, middle temporal gyrus; Orb. S, orbital sulcus; OTS, occipito-temporal sulcus; PAC, primary auditory cortex; pITS, posterior inferotemporal sulcus; pMTG, posterior middle temporal gyrus; PoCeS, postcentral sulcus; POS, parieto-occipital sulcus; SFS, superior frontal sulcus; SPRCeS, superior precentral sulcus; STG, superior temporal gyrus; STS, superior temporal sulcus; TOS, transverse occipital sulcus.

selective for tools as opposed to animal vocalizations, although the interpretation of these negative BOLD signals remains unclear. The other seven ROIs showed positive responses to both tool and animal sounds relative to silence but were significantly

preferential for one or the other sound category. This included the bilateral mSTG foci for animal sounds and the bilateral pLaS, left mIFS, left VPMC, and left IPoCeC for tool sounds. Together, these ROI data demonstrated that a range of differential re-

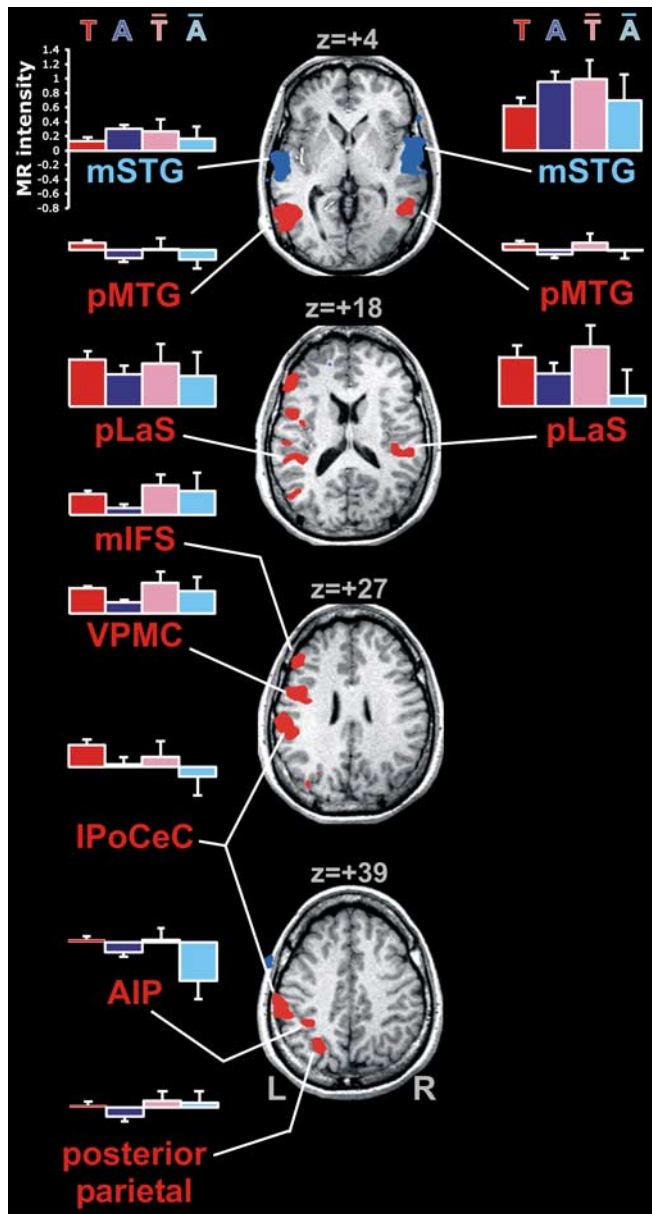


Figure 2. Group-averaged ROI analysis of activated areas from Figure 1*d*. Charts illustrate the mean and SE across 20 participants (unthresholded) of the relative BOLD signal intensity (arbitrary units) in response to correctly categorized tool sounds (“T,” red) versus silence and animal sounds (“A,” dark blue) versus silence. They also depict responses to animal sounds miscategorized as tools (“ \bar{T} ,” pink) and tool sounds miscategorized as animals (“ \bar{A} ,” light blue), across 17 participants who made errors in both sound categories. The left mesial frontal focus from Figure 1*e* showed BOLD responses below baseline (silence) to all sounds (data not shown). Brain images (axial slices) are from one participant and transformed into Talairach coordinate space.

sponses were evoked by listening to and correctly categorizing tool versus animal sounds.

We additionally charted brain responses in the ROIs when sounds were incorrectly categorized, including animal sounds miscategorized as tools (Fig. 2, “ \bar{T} ,” pink) and tool sounds miscategorized as animals (“ \bar{A} ,” light blue), both relative to silence. The mSTG foci were more strongly activated by animal vocalization sounds (middle two columns) than tool sounds, regardless of whether or not they were correctly categorized. These findings were consistent with the placement of the mSTG at early to intermediate auditory processing stages as opposed to more high-level conceptual stages, as is addressed further below.

In striking contrast, all of the ROIs that were preferential for correctly categorized tool sounds relative to animal sounds (red > dark blue) were also preferentially activated by animal sounds judged to be tool sounds relative to tool sounds judged to be animal sounds (pink > light blue). The miscategorization analysis revealed a data trend that clearly showed the “tool-related” network (red foci) being preferentially activated when sounds, whether correctly or incorrectly categorized, were perceived as being produced by tools.

Manipulating virtual tools

One possible explanation for the strong left-lateralized activation evoked by the tool-related sounds (red) was that these regions may be associating dominant right-hand and arm motor actions typically correlated with tool use. To explore this possibility, 12 of the 20 participants additionally performed a separate motor task, usually during the same scanning session. In a block paradigm, they alternated between making right-hand and arm movements as if manipulating a variety of virtual tools (simulating hammering, sanding, sawing, ratcheting, etc.) versus resting the hand motionless. No sound stimuli were presented during these scans except for tone pips to cue when the task periods began and ended.

Making hand manipulations as if using tools evoked the strongest activation (Fig. 3, green) in the hand and arm representations of the primary motor and somatosensory cortices (“M1” and “S1,” respectively) of the left hemisphere and in the right cerebellar cortex (data not shown). Other cortical regions were also significantly activated by manipulating virtual tools and were essentially consistent with previous studies involving actual tool or object manipulations in contrast to other movements (Binkofski et al., 1999; Moll et al., 2000; Amedi et al., 2001; Choi et al., 2001). However, activity in the left and right pMTG regions was more pronounced in, if not unique to, the present study. This may have been attributable to the participants imagining the sounds and/or visualizing the pantomimed tools, in contrast to the motionless rest condition. Thus, some of the activity (green) may reflect visualizations or other forms of mental imagery in addition to, or instead of, overt motor processing. The intermediate colors yellow and cyan depict regions of overlap with the tool versus animal sound recognition data, respectively.

The main result of the virtual tool manipulation experiment was that most of the cortical foci showing preferential activity to hearing and categorizing tool-related sounds (as opposed to animal sounds) were also activated, at least in part, by actual right-hand movements mimicking the use of tools in the absence of sound (Fig. 3, yellow). This included the left mIFS, left VPMC, left IPoCeC, left AIP, bilateral pLaS, and bilateral pMTG foci. The overlap between the tool sound processing foci and the hand manipulation foci was also evident on a voxelwise basis in most individuals.

Spectral and phase analysis of sound signals

Some of the differences in cortical activity evoked by categorizing tool versus animal sounds described above appeared to be attributable, in part, to relatively high-level associative processing with other sensory modalities (multimodal or supramodal processing), notably including motor-related cortices. However, some of the differential brain activation may have been attributable to differences in the acoustic signal properties (low-level auditory processing) between these two conceptual categories of sound (Nelken et al., 1999; Reide et al., 2001; Ehret and Riecke, 2002; Lewicki, 2002). To explore this possibility, we quantitatively compared (Fig. 4) the acoustical features of the tool and animal sound stimuli we presented. Additionally, because the left and

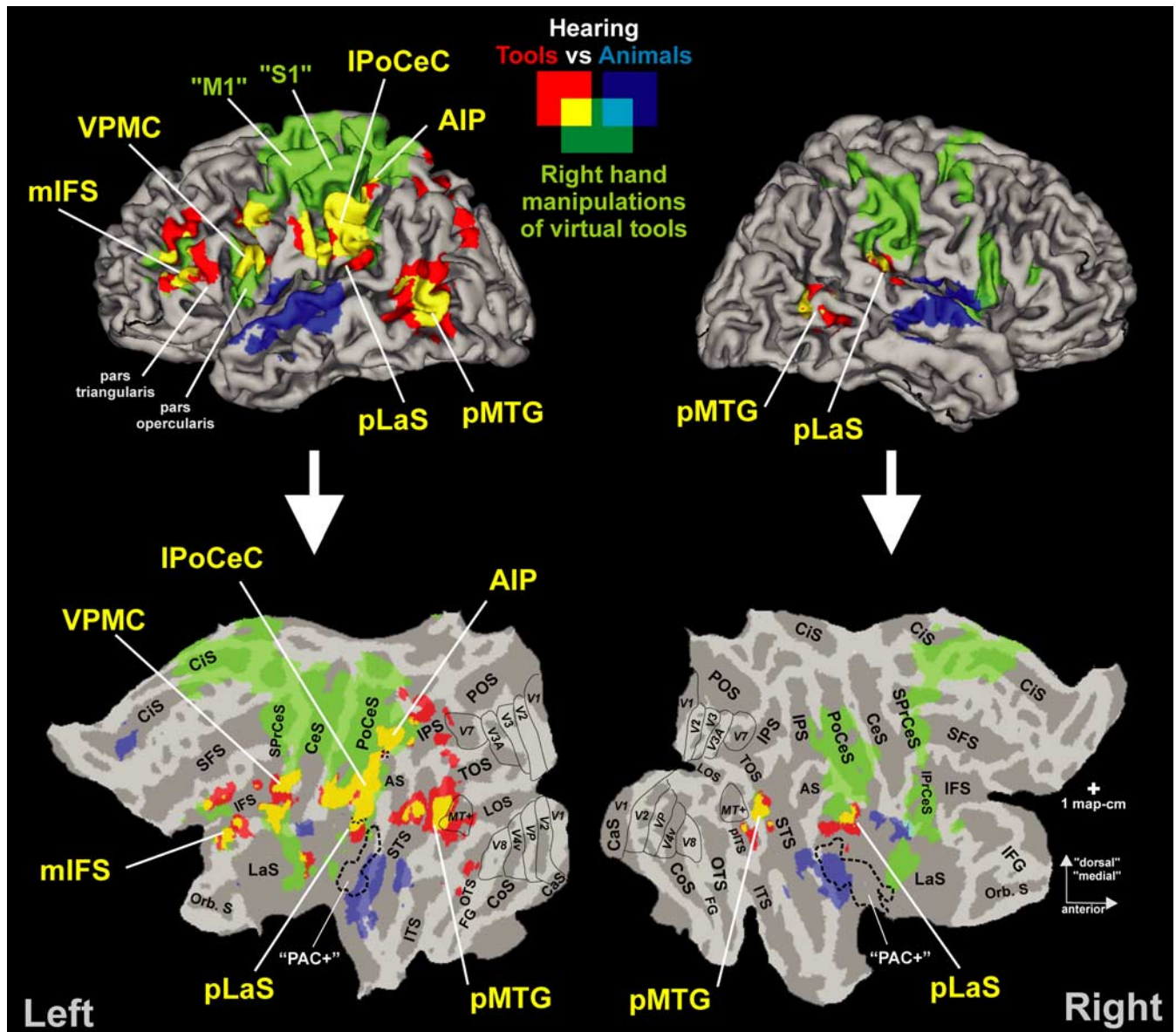


Figure 3. Group activation data evoked by manipulations of virtual tools using the right hand (green; $n = 12$; corrected $\alpha < 0.05$) superimposed onto the data from Figure 1, *d* and *e*. The intermediate colors yellow and cyan depict regions of overlap. Some of the cerebellar cortex activation extended into the ventral occipital lobes but was removed for clarity. Other conventions are as in Figure 1.

right mSTG foci for animal sounds directly overlapped cortex found previously to be preferentially activated by speech versus nonspeech sounds (Belin et al., 2000; Binder et al., 2000; Fecteau et al., 2004; Lewis et al., 2004), we also analyzed a wide variety of human speech sounds, expressions, and utterances for quantitative comparison. This included samples from American English dialogs and monologs, which were processed (and matched in duration and loudness) using exactly the same techniques as for the tool and animal sounds.

Spectrographs (Fig. 4*a–c*) illustrate the energy of specific frequencies present over the duration of each sound. To quantify some of these complex signal attributes, we first examined the overall power spectra from 0 to 20 kHz (shown to the right of each spectrograph). On average (Fig. 4*d*), the animal and speech sounds (vocalizations) contained greater similarity in their overall normalized power spectra, including relatively greater power in the ~500–3500 Hz range. However, the power spectrum did not capture or characterize any of the potential interactions or relationships be-

tween frequency components in the acoustic signals or the evolution of spectral features over time. Presumably, some of these more complex spectral and temporal relationships could distinguish the different categories of sound (Nelken et al., 1999; Reide et al., 2001; Ehret and Riecke, 2002). Thus, we explored two advanced signal processing techniques in an attempt to distinguish these three conceptually distinct categories of sound.

Because vocalization sounds are known to have a large degree of harmonic content (Reide et al., 2001; Ehret and Riecke, 2002), we examined the HNR of all of the sounds. This measure compared the acoustic energy of the harmonic components (periodic signals) over time with that of noise, in which noise was defined as the remaining nonharmonic, irregular, or chaotic acoustic energy in the signal. The HNR values correlated well with the spectrographs. For instance, the spectrographs of the animal sound (Fig. 4*b*; rooster call) and speech sound (Fig. 4*c*; woman answering “hello” on a telephone) both showed clear bands of energy in frequency (dark horizontal banding patterns), which were consistent with their having large HNR values

or “harmonic content.” In contrast, the spectrograph of the tool sound (Fig. 4*a*; a saw cutting wood with an up, down, up, down stroke cycle) showed a broad and nearly continuous range of frequencies within each stroke and consequently had a very low (negative) HNR value. On average (Fig. 4*e*), the speech sounds had greater HNR values than the animal sounds, and both the animal and human sounds (vocalizations) had much greater harmonic content than the tool sounds.

Interestingly, the tool sounds that were miscategorized as animal sounds contained relatively higher HNR values (average, -0.2 dB; 21 sounds) than the correctly categorized tools sounds (-1.1 dB), and the miscategorized animal sounds had lower HNR values ($+6.1$; 24 sounds) than the correctly categorized animal sounds ($+9.9$ dB). This error trend supported the notion that HNR signal content (or lack thereof) may have influenced the perception of the sounds.

To examine temporal dynamics of the sounds, we adopted a bicoherence analysis (Fig. 4*f*) that quantified the degree of phase coupling of frequencies within the different sounds (Sigl and Chamoun, 1994). Nerve fibers in the cochlea can respond to pure sine-wave tones by synchronizing (spiking) with a particular phase angle of the sound (Warren, 1999). Such “phase locking” can occur with frequencies up to ~ 4000 Hz, thereby serving as one means for the brain to encode or neurally represent tones. Phase information is known to be used by the auditory system, such as when hearing binaural beats (akin to hearing beats when tuning guitar strings) and for spatial hearing (localizing a sound based on phase angle differences at each ear) (Blauert, 1997). The bicoherence analysis, unlike the power spectra, quantified the degree of phase coupling (monaurally; in only the left ear channel of a given sound) that existed between different frequency pairs independent of amplitude. Frequency components that were coupled in-phase (having a nonrandom relationship) produced a greater bicoherence value, which is reflected in the one-dimensional profiles illustrated. The bicoherence analysis revealed a greater degree of similarity between the animal and human vocalizations relative to the tool sounds, especially in the 250–1400 Hz range. Whether phase-coupling attributes can significantly influence sound recognition processes remains to be explored but appear to be a potential cue for distinguishing vocalizations from tool-related sounds. Together, the above quantitative analyses suggest that the signal attributes of the animal versus tool sounds could have accounted for some of the differential fMRI activation, quite likely including the activity in the bilateral mSTG foci.

Discussion

Animal sounds and the bilateral mSTG

Together with previous fMRI studies, the present data suggest that the bilateral mSTG foci, preferential for animal sounds, represent “intermediate” or pre-representational sound processing

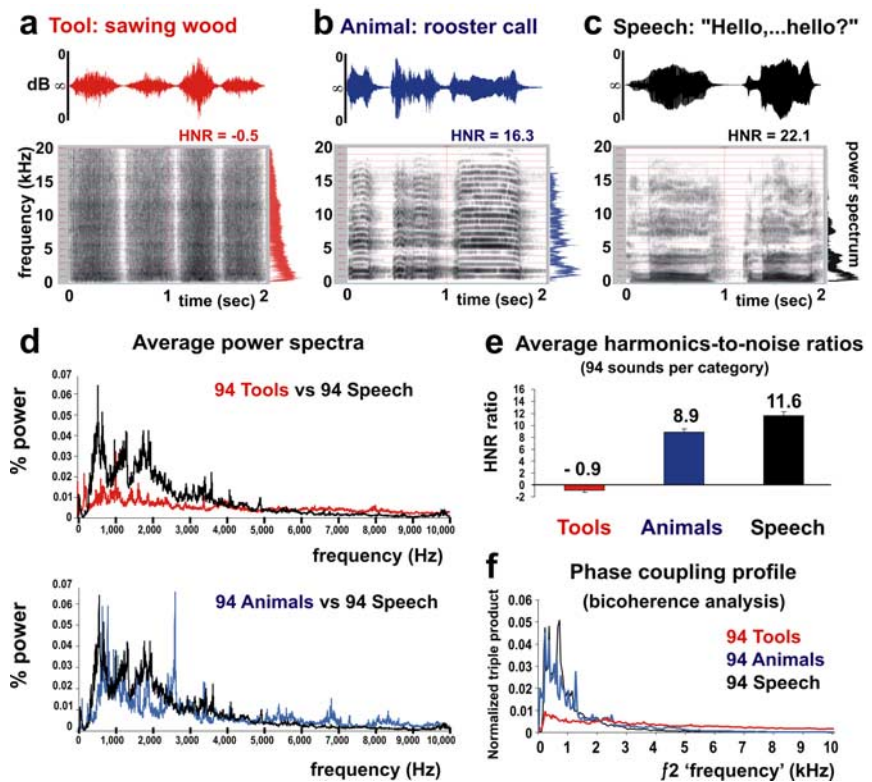


Figure 4. Quantitative comparisons of acoustical differences between categories of sound, illustrating the similarity between animal sounds (blue) and human speech sounds (black) relative to tool sounds (red). *a–c*, Example sound stimuli illustrating the corresponding amplitude plot, spectrograph, power spectrum (percentage power vs frequency), and HNR. *d*, Power spectra averaged across the 94 retained tool sounds (top) and 94 animal sounds (bottom) compared with 94 human speech sounds. Note that, although the average RMS power was balanced across sound categories, the average power showed differences. The average power beyond 10 kHz in all categories was negligible and thus is not shown. *e*, Mean (plus variance) of HNR values across sound categories. *f*, Bicoherence analysis comparing the three categories of sound, illustrating different phase-coupling profiles. These and other samples of the sound stimuli can be heard at www.jneurosci.org, as supplemental material.

stages of a cortical sound recognition system. The bilateral mSTG foci directly overlapped a progression of cortex reported to be more responsive to passively heard spoken words versus tones, and tones versus white noise (Binder et al., 2000; Wessinger et al., 2001; Lewis et al., 2004). This progression, extending multidirectionally from the primary auditory cortex out to and including the mSTG, appears to reflect a hierarchy of increasing responsiveness to increasing acoustic structure. Additionally, previous neuroimaging and electrophysiological studies have reported preferential activity to human voices or vocalizations along cortex overlapping (or near) the bilateral mSTG, in contrast to a variety of complex control sounds, including scrambled voices or musical instruments (Belin et al., 2000; Levy et al., 2001; Fecteau et al., 2004). Our ROI analysis indicated that the animal sounds (Fig. 2, dark blue and pink), in contrast to the tool sounds, produced greater activation in the mSTG foci regardless of whether the sounds were correctly categorized. Similarly, these foci were found to be relatively insensitive to environmental sound recognition per se (Lewis et al., 2004). Thus, the mSTG foci appear to represent stages that are primarily before high-level semantic, lexical, or other representational processing (Damasio et al., 1996; Binder et al., 1997; Näätänen and Winkler, 1999; Tranel et al., 2003).

The above progression of auditory cortex shares both anatomical and functional homology with some of the macaque lateral and anterior “belt” and “parabelt” auditory areas that surround primary auditory cortex (Rauschecker, 1998; Kaas et al., 1999). These areas include increasing proportions of neurons selective

Table 1. Tool versus animal sound processing foci

Side	Anatomical location	Brodmann area	Talairach coordinates			Volume (mm ³)
			x	y	z	
Tools > animals						
R	pMTG (STS)	Areas 37, (39, 22)	49	−51	5	1853
L	pMTG (STS)	Area 37, (19)	−51	−57	3	10,408
R	pLaS	Areas 13, 40	48	−29	20	1254
L	pLaS ^a	Areas 13, 40	−44	−35	18	477
L	IPoCeC ^a (TPJ)	Area 40	−55	−28	33	8438
L	AIP ^a	Areas 40 (2)	−39	−40	42	745
L	Parietal (multiple foci)	Precuneus	−28	−64	40	2327
L	mIFS	Areas 46, 45	−49	31	21	3505
L	VPMC (area 6)	Areas 6, 44	−46	3	27	4141
Animals > tools						
R	mSTG	Area 22	56	−8	5	6191
L	mSTG	Area 22	−60	−15	2	6190
L	Mesial frontal gyrus	Areas 10, 9	−10	51	11	850

Talairach coordinates of cortical foci preferentially activated by tool or animal sounds across 20 participants. The AIP boundary (Fig. 1e, black dotted line) was established at slightly greater threshold settings. The dorsal boundary of the left pLaS (dark gray dotted line) was defined by the dorsalmost extent (at $z = +25$) of the right pLaS focus in the volume data, which roughly separated the inferior parietal lobule from the insular cortex.

^aThe left hemisphere pLaS and IPoCeC and AIP ROIs were one contiguous focus at $\alpha < 0.05$.

for components of species-specific monkey calls. Thus, they are thought to subserve the pre-processing of vocalization and communication sounds (Rauschecker, 1998).

We hypothesize that the preferential activity in the mSTG for animal sounds reflects the preferential processing of certain signal attributes, such as harmonic or phase-coupling content (Fig. 4), which are characteristic of vocalization sounds and tend to be less pronounced or less consistent in tool-related sounds. Because most human listeners are exceedingly experienced with processing and recognizing human speech (from birth and perhaps *in utero*), the bilateral mSTG foci may represent stages that are, or can become, optimally organized to process the prevailing signal components of vocalizations (predominantly human speech but also animal vocalizations). Thus, the mSTG regions may represent rudimentary stages of a sound recognition system that are optimized for spoken language processing.

Tool sounds and action knowledge

Right-handed people typically learn of characteristic tool sounds while manipulating a tool with their right-hand and viewing the complex motions of the hand and tool. The cortex that was preferential for processing tool sounds (Fig. 3, yellow and red) may be involved in associating or matching motor ideas regarding right arm and hand manipulations (and possibly visualizing the manipulations) that could have been correlated with sound production. Ostensibly, the collection of left-lateralized regions preferential for tool sounds was consistent with representing part of a “mirror system,” as reported in both human (Iacoboni et al., 1999; Rizzolatti et al., 2002; Grèzes et al., 2003) and monkey (Kohler et al., 2002). In particular, the left VPMC and/or mIFS foci may share homology with the macaque area F5, which contains “audiovisual mirror neurons” that discharge when the animal performs a specific action, views the related action, or hears the related sound in isolation (Kohler et al., 2002). Similarly, portions of the left IPoCeC may share homology with the macaque area 7b, which is thought to be involved in the perception of the space within grasping range and in the organization of contralateral arm movements toward stimuli presented in that space (Rushworth et al., 2001; Burton, 2002).

The bilateral pLaS foci may share homology with auditory-somatosensory convergence sites in the retroinsular cortex of the macaque (Hackett et al., 2001; Schroeder et al., 2001; Schroeder and Foxe, 2002; Fu et al., 2003). This includes the caudomedial auditory belt area, which in humans is a region maximally activated, for instance, by the combined presentation of sandpaper-like sounds

together with feeling sandpaper being rolled over the right hand (Foxe et al., 2002). The pLaS foci may also overlap the somatosensory area S2, which has a role in active hand manipulations and integrated sequential touching, such as for identifying objects based on touch (Burton, 2002; Grèzes et al., 2003). Together, the preferential activation to tool sounds in the left VPMC and mIFS, IPoCeC, AIP, and bilateral pLaS was consistent with the mirroring of motor actions that were likely to have led to the tool-related sounds.

The left and right pMTG activation for tool sounds may have reflected the processing of dynamic motion associations (motor or visual) or more abstract “action knowledge” representations associated with the sounds (Martin et al., 1996; Grossman et al., 2002; Phillips et al., 2002; Kellenbach et al., 2003). In our previous study, these foci were activated by a much wider range of recognizable environmental sounds (depicting actions) than just tool-related sounds (Lewis et al., 2004). Moreover, pMTG activity has been reported in response to viewing tools or objects in motion (Beauchamp et al., 2002, 2004; Manthey et al., 2003) or pictures of tools associated with action (Chao et al., 1999; Grèzes and Decety, 2002). Thus, the bilateral pMTG activity may be reflecting some form of visual or mental imagery of the motion dynamics associated with the sound.

Together, the network of regions preferential for tool sounds may serve to match or associate multisensory knowledge to reason as to how the tool sounds might have been produced, revealing a high-level auditory “how” pathway for purposes of recognition (Johnson-Frey, 2003; Kellenbach et al., 2003). Thus, what we perceive when we hear tool sounds may not be determined so much by the signal characteristics themselves but rather by the relationship between the sound and our experiences with the probable actions that produced them (Schwartz et al., 2003; Körding and Wolpert, 2004).

From multisensory perception to conception and language

Strikingly, the left pMTG focus for tool sounds and the left mSTG focus for animal sounds (Fig. 1d) complemented, by partially overlapping, the reported large-scale cortical architecture for tool versus animal word-form (lexical/semantic) knowledge (Damasio et al., 1996). In particular, lesions to the left mSTG region (and cortex farther ventral) can lead to deficits in naming pictures of animals, whereas lesions to the left pMTG produce deficits more specific to naming tools. These regions may thus represent early stages of a language-specific pathway that leads to the verbal representation and identification of the perceived sounds (Tranel et al., 2003). However, why would word-form

knowledge for the category of tools be represented in the posterior temporal lobe and that for animals be situated farther anterior, and not, say, vice versa?

The present data are suggestive of a multimodal sensory-driven mechanism that may explain this large-scale architecture. We speculate that tool sounds typically have a greater degree of intermodal invariant associations with the other sensory modalities than do most vocalization sounds (Lewkowicz, 2000; Calvert and Lewis, 2004). For instance, pounding with a hammer will typically expose the perceiver to strongly correlated temporal information from the auditory, motor, tactile, and visual modalities. The pMTG represents cortex that is well situated, for compactness of cortical “wiring,” to receive and associate multiple sensory inputs, because they are approximately located between the primary sensory domains for hearing, motor/touch, and vision (Van Essen, 1997; Amedi et al., 2001; Beauchamp et al., 2002, 2004; Lewis et al., 2004).

In contrast, vocalization sounds may evoke a relatively greater degree of processing specific to the auditory modality, such as extracting harmonic or phase-correlated information within the sound (e.g., leading to phoneme and word recognition). Such processing might ideally be situated closer to the primary and nearby auditory cortices (such as the mSTG foci), again for compactness of cortical wiring, because there would be relatively fewer dynamic multimodal associations to be made. A notable exception would include lip reading while listening to speech, which involves audiovisual integration that can greatly aid in speech perception. Although lip reading can activate the mSTG foci (Calvert et al., 1997), the regions most sensitive to the integration of audio and visual speech overlap cortex near the pMTG foci (Calvert et al., 2000; Calvert and Lewis, 2004). These findings support the proposed role of the pMTG in processing dynamic multimodal “actions,” which more typically may apply to tool

sounds, but in some circumstances would apply to observing (hearing and seeing) the production of vocalizations with correlated mouth, face, and body gesticulations.

Curiously, there was a moderate degree of overlap between the classically defined Broca’s area for language production (Fig. 3, compare near pars opercularis and pars triangularis) with the left VPMC and mIFS foci, which were preferential for tool sounds and virtual tool manipulations. Some theories of the evolution of speech perception (e.g., the “motor theory”) posit that the auditory representation of speech sounds would ideally be situated near motor cortices (Liberman and Mattingly, 1985; Iacoboni et al., 1999; Holden, 2004). In particular, speech production and other forms of communication are primarily learned through mimicry, involving the sequential coordination of muscle movements of the vocal chords and/or the arms and hands. Aspects of tool use are also learned through mimicry, and the resulting sounds (their “meaning”) may thus be associated with such motor actions. Thus, the present data appear to lend support to a motor theory behind the evolution of communication and language in humans, reflecting a possible link between sound recognition, imitation, gesturing, and language articulation.

Together, the present findings are consistent with the existence of sensory- and multisensory-driven mechanisms that may establish the gross cortical organization of the human auditory system, wherein animal vocalization sounds evoke greater processing demands on the mSTG bilaterally, whereas tool sounds impose greater demands on mostly left hemisphere audiomotor association cortices and the bilateral pMTG (possibly audiovisual association cortices). This organization, via multisensory “bootstrapping,” may then lead to the reported large-scale cortical architecture for word-form knowledge, including the distinct conceptual categories of man-made artifacts (tools) and living things (animals).

Appendix

Table A. Tool sounds

Adding machine, manual (20)	Ice scraper, scraping windshield (20)	Scissors, cutting hair (17)
Air wrench, putting on lug nuts (19)	Igniter, handheld, lighting arc welder (19)	Scissors, cutting paper (20)
Bench vice, cranking handle (19)	Knife, chopping carrots (20)	Scissors, cutting plastic (20)
Broom, short bristled, light sweeping (18)	Knife, cutting cardboard (18)	Scissors, cutting wallpaper and snipping (17)
Broom, sweeping and dapping (20)	Knife, sharpening with fast motion (20)	Scraping paint off wall (20)
Brush, stiff bristles sweeping (13)	Knife, sliding into sheath (20)	Screwdriver, screwing into wood (19)
Caulking gun, fast application (20)	Letter opener, cutting envelope open (20)	Self-inking stamp, manual (20)
Caulking gun, slow application (20)	Metal filing, fast (20)	Shaking up can of spray paint (20)
Chamois cloth, washing window (2)	Mixer #1, metallic sound (20)	Sharpening chisel, high pitch (19)
Chisel into wood, fast grating (20)	Mixer #2, beating eggs in bowl (19)	Sharpening chisel, low pitch (20)
Chisel into wood, slow grating (20)	Mortar and pestle, grinding (20)	Sickle, slashing hay (20)
Coping saw (20)	Nail gun, pneumatic (20)	Spade, chopping action (20)
Filing metal, slowly (20)	Padlock, tinkering with key and opening (18)	Spanner, nut turning on screw (20)
Filing wood, fast, bidirectional (20)	Paint brush, fast application of paint (19)	Spray bottle, slow high-pitched sprays (19)
Filing wood, one direction (20)	Paint roller, wet sound (20)	Spray bottle, spraying liquid quickly (16)
Glass cutter (20)	Paper cutter, cutting (20)	Staple remover, fidgeting sounds (19)
Hack saw, copper pipe (20)	Paper punch, low pitch (20)	Stapler being opening (20)
Hammer removing nail #1, hard (12)	Plunger, freeing clogged drain (13)	Stapler, heavy stapling (19)
Hammer removing nail #2, light (14)	Pump ratchet (19)	Stapler, light stapling (20)
Hammer, hitting metal pipes (20)	Rake, raking gravel (20)	Stapling gun (20)
Hammer, tack, light pounding (19)	Rake, raking leaves (20)	Tape dispenser, tape pulled and torn (17)
Hammering chisel into concrete (20)	Rasp, filing wood (20)	Tin snips, cutting metal (19)
Hammering chisel into wood (20)	Ratchet screwdriver, fast high pitch (19)	Toilet plunger, water sloshing (15)
Hammering horseshoe on anvil (20)	Ratchet, high pitch, metallic (19)	Trowel, digging and scraping sand (20)
Hammering metal sheet (20)	Ratchet, slow low-pitch turning (19)	Trowel, plastering wall (19)
Hammering nail, into wall, with echoes (20)	Rubber stamp, stamping (20)	Typewriter, fast carriage rolling (19)
Hammering nail, into wood (20)	Sandpapering, fast (19)	Whiplashes (20)
Hand plane, slow draw, tinny sound (9)	Sandpapering, slow (20)	Wire brush, cleaning metal (20)
Hand vice grips, closing and opening (20)	Sawing wood, drawn back once slowly (2)	Wire brush, removing paint (20)
Hatchet, chopping wood (20)	Sawing wood, fast draws (20)	Wire brush, scrubbing floor (19)
Ice pick, picking (20)	Sawing wood, slow draws (19)	Writing with ballpoint pen (15)
		Writing with pencil (19)

Table B. Animal sounds

Alligator, adult growl #1 (20)	Bird, whippoorwill, call (20)	Lion, snap and growl (20)
Alligator, adult growl #2 (20)	Bird, wren (20)	Monkey, howler (15)
Alligator, baby calling (6)	Bird, yellowthroat bird (20)	Monkey, pygmy marmoset squeal (20)
Bear, growl (18)	Bull, bellowing (20)	Monkey, small animal screech (20)
Bear, large, breathing (3)	Bull, howling (20)	Panda cub, mewling (20)
Bear, large, chewing and licking (20)	Camel, groan (19)	Penguin, baby emperor, talking (13)
Bear, large, growl and roar (20)	Cat, meowing (12)	Penguin, single short call (19)
Bear, large, snorting (20)	Cat, noisy growling (20)	Pig, deep grunting (20)
Bird, blackbird chirp (18)	Cat, purring (17)	Pig, grunt and snort (20)
Bird, chickadee, single call (20)	Dog, crying-yowl (20)	Pig, squeal (18)
Bird, chicken, clucking and call (20)	Dog, howl startup (19)	Rhinoceros, grunt and whine (18)
Bird, cockatoo screeching (20)	Dog, whimpering (20)	Rhinoceros, inhale and groan (20)
Bird, cockatoo calling (20)	Dolphin, chirp and high pitch (20)	Sea lion, barking (18)
Bird, cuckoo (18)	Dolphin, chirps (18)	Sea lion, distant higher pitch (19)
Bird, dove cooing (20)	Dolphin, screeches (20)	Sea lion, single roar (19)
Bird, duck quacking (20)	Donkey, braying (19)	Sea lion, single wavering call (16)
Bird, eagle calling (20)	Donkey, ramp up sound (20)	Sea otter, squealing (20)
Bird, falcon, raspy call (15)	Elephant, cry (19)	Seal, elephant, call (16)
Bird, finch song (20)	Elephant, groan (11)	Seal, single bark (20)
Bird, heron, screech noises (20)	Elephant, groan-sigh (16)	Tiger, growl (19)
Bird, killdeer calling (19)	Elephant, trumpets (20)	Turkey, gobbling (20)
Bird, loon, single call (19)	Elk, screeching call (20)	Walrus, female, single roar (19)
Bird, macaw, raspy call (20)	Frog, ribbiting (20)	Walrus, male, single roar (16)
Bird, mockingbird, call (20)	Gorilla, vocalizing (20)	Whale, beluga (20)
Bird, nightingale, shrill (19)	Horse, snorts (20)	Whale, humpback, moans (16)
Bird, parrot (19)	Horse, whinny (20)	Whale, killer whale, call (19)
Bird, peacock (20)	Insect, cricket chirping (20)	Whale, pilot (19)
Bird, rooster call (20)	Jaguar, growl/purr (19)	Whale, pilot, high-pitch call (20)
Bird, skylark, complex call (20)	Jaguar, snap and growl (19)	Whale, pilot, spitting noises (14)
Bird, sparrow, complex trills (20)	Leopard, growl (18)	Wildebeest, bleating (20)
Bird, sparrow, mellow call (20)	Leopard, yowl (19)	Wildebeest, moans (soft, low pitch) (19)
		Zebra, barking (19)

The 94 animal sounds that were presented and retained for analyses are listed. Parentheses indicate the number of participants for which a given sound was correctly categorized.

Please visit www.jneurosci.org to view on-line supplemental material.

References

- Amedi A, Malach R, Hendler T, Peled S, Zohary E (2001) Visuo-haptic object-related activation in the ventral visual pathway. *Nat Neurosci* 4:324–330.
- Bandettini PA, Jesmanowicz A, Wong EC, Hyde JS (1993) Processing strategies for functional MRI of the human brain. *Magn Reson Med* 30:161–173.
- Beauchamp M, Lee K, Haxby J, Martin A (2002) Parallel visual motion processing streams for manipulable objects and human movements. *Neuron* 34:149–159.
- Beauchamp MS, Lee KM, Argall BD, Martin A (2004) Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron* 41:809–823.
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B (2000) Voice-selective areas in human auditory cortex. *Nature* 403:309–312.
- Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T (1997) Human brain language areas identified by functional magnetic resonance imaging. *J Neurosci* 17:353–362.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Rao SM, Cox RW (1999) Conceptual processing during the conscious resting state: a functional MRI study. *J Cogn Neurosci* 11:80–95.
- Binder JR, Frost J, Hammeke T, Bellgowan P, Springer J, Kaufman J, Possing E (2000) Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex* 10:512–528.
- Binkofski F, Buccino G, Posse S, Seitz RJ, Rizzolatti G, Freund H-J (1999) A fronto-parietal circuit for object manipulation in man: Evidence from an fMRI-study. *Eur J Neurosci* 11:3276–3286.
- Blauert J (1997) Spatial hearing with one sound source. In: *Spatial hearing*, Chap 2. Cambridge, MA: MIT.
- Boersma P (1993) Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proc Inst Phon Sci* 17:97–110.
- Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, Seitz RJ, Zilles K, Rizzolatti G, Freund H-J (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur J Neurosci* 13:400–404.
- Burton H (2002) Cerebral cortical regions devoted to the somatosensory system: results from brain imaging studies in humans. In: *The somatosensory system: deciphering the brain's own body image* (Randall J, Nelson PD, ed), pp 27–72. New York: CRC.
- Calvert G, Bullmore ET, Brammer MJ, Campbell R, Williams SCR, McGuire PK, Woodruff PWR, Eversen SD, Anthony S (1997) Activation of auditory cortex during silent lip reading. *Science* 276:593–596.
- Calvert GA, Lewis JW (2004) Hemodynamic studies of audio-visual interactions. In: *Handbook of multisensory processing* (Calvert GA, Spence C, Stein B, eds), pp 483–502. Cambridge, MA: MIT.
- Calvert GA, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr Biol* 10:649–657.
- Chao LL, Martin A (2000) Representation of manipulable man-made objects in the dorsal stream. *NeuroImage* 12:478–484.
- Chao LL, Haxby JV, Martin A (1999) Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat Neurosci* 2:913–919.
- Choi SH, Na DL, Kang E, Lee KM, Lee SW, Na DG (2001) Functional magnetic resonance imaging during pantomiming tool-use gestures. *Exp Brain Res* 139:311–317.
- Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29:162–173.
- Damasio H, Grabowski TJ, Tranel D, Hichwa RD, Damasio RD (1996) A neural basis for lexical retrieval. *Nature* 380:499–505.
- De Renzi E, Lucchelli F (1988) Ideational apraxia. *Brain* 111:1173–1185.
- Ehret G, Riecke S (2002) Mice and humans perceive multiharmonic communication sounds in the same way. *Proc Natl Acad Sci USA* 99:479–482.
- Fecteau S, Armony JL, Joannette Y, Belin P (2004) Is voice processing species-specific in human auditory cortex? An fMRI study. *NeuroImage* 23:840–848.
- Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM (2002) Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J Neurophysiol* 88:540–543.
- Fu K-MG, Johnston TA, Shah AS, Arnold L, Smiley J, Hackett TA, Garraghty

- PE, Schroeder CE (2003) Auditory cortical neurons respond to somatosensory stimulation. *J Neurosci* 23:7510–7515.
- Gainotti G, Silveri M, Daniele A, Giustolisi L (1995) Neuroanatomical correlates of category-specific impairments: a critical survey. *Memory* 3:247–264.
- Grèzes J, Decety J (2002) Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia* 40:212–222.
- Grèzes J, Armony JL, Rowe J, Passingham RE (2003) Activations related to “mirror” and “canonical” neurones in the human brain: an fMRI study. *NeuroImage*:928–937.
- Grossman M, Koenig P, DeVita C, Glosser G, Alsop D, Detre J, Gee J (2002) The neural basis for category-specific knowledge: an fMRI study. *NeuroImage* 15:936–948.
- Hackett TA, Preuss TM, Kaas JH (2001) Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J Comp Neurol* 441:197–222.
- Holden C (2004) The origin of speech. *Science* 303:1316–1319.
- Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G (1999) Cortical mechanisms of human imitation. *Science* 286:2526–2528.
- Ilmberger J, Rau S, Noachtar S, Arnold S, Winkler P (2002) Naming tools and animals: asymmetries observed during direct electrical cortical stimulation. *Neuropsychologia* 40:695–700.
- Johnson-Frey SH (2003) What’s so special about human tool use? *Neuron* 39:201–204.
- Kaas JH, Hackett TA, Tramo MJ (1999) Auditory processing in primate cerebral cortex. *Curr Opin Neurobiol* 9:164–170.
- Kellenbach ML, Brett M, Patterson K (2003) Actions speak louder than functions: the importance of manipulability and action in tool representation. *J Cogn Neurosci* 15:30–46.
- Kéri S (2003) The cognitive neuroscience of category learning. *Brain Res Rev* 43:85–109.
- Kohler E, Keysers C, Umiltà A, Fogassi L, Gallese V, Rizzolatti G (2002) Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297:846–848.
- Komatsu LK (1992) Recent views of conceptual structure. *Psychol Bull* 112:500–526.
- Körding KP, Wolpert DM (2004) Bayesian integration in sensorimotor learning. *Nature* 427:244–247.
- Levy DA, Granot R, Bentin S (2001) Processing specificity for human voice stimuli: electrophysiological evidence. *NeuroReport* 28:2653–2657.
- Lewicki MS (2002) Efficient coding of natural sounds. *Nat Neurosci* 5:356–363.
- Lewis JW, DeYoe EA (2003) Animal vs tool sounds evoke different activation patterns in right vs left handed subjects. *Hum Brain Mapp Abstr* 19:1488.
- Lewis JW, Beauchamp MS, DeYoe EA (2000) A comparison of visual and auditory motion processing in human cerebral cortex. *Cereb Cortex* 10:873–888.
- Lewis JW, Wightman FL, Breczynski JA, Phinney RE, Binder JR, DeYoe EA (2004) Human brain regions involved in recognizing environmental sounds. *Cereb Cortex* 14:1008–1021.
- Lewkowicz DJ (2000) The development of intersensory temporal perception: an epigenetic systems/limitations view. *Psychol Bull* 126:281–308.
- Lieberman AM, Mattingly IG (1985) The motor theory of speech perception revised. *Cognition* 21:1–36.
- Lusthaus D (2002) Buddhist phenomenology: a philosophical investigation of yogaacaara Buddhism and the Ch’eng wei-shih lun, Chap 2. London: Curzon.
- Manthey S, Shubotz RI, von Cramon DY (2003) Premotor cortex in observing erroneous action: an fMRI study. *Cognit Brain Res* 15:296–307.
- Martin A (2001) Functional neuroimaging of semantic memory. In: *Handbook of functional neuroimaging of cognition* (Cabeza R, Kingstone A, eds), pp 153–186. Cambridge, MA: MIT.
- Martin A, Wiggs CL, Ungerleider LG, Haxby JV (1996) Neural correlates of category-specific knowledge. *Nature* 379:649–652.
- Medin DL, Coley JD (1998) Concepts and categorization. In: *Perception and cognition at century’s end: history, philosophy, theory* (Handbook of perception and cognition) (Hochberg J, Cutting JE, eds), pp 403–432. New York: Academic.
- Moll J, de Oliveira-Sousa R, Passman LJ, Cunha FC, Souza-Lima F, Andreiuolo PA (2000) Functional MRI correlates of real and imagined tool-use pantomimes. *Neurology* 54:1331–1336.
- Moore CJ, Price CJ (1999) A functional neuroimaging study of the variables that generate category-specific object processing differences. *Brain* 122:943–962.
- Näätänen R, Winkler I (1999) The concept of auditory stimulus representation in cognitive neuroscience. *Psychol Bull* 125:826–859.
- Nelken I, Rotman Y, Yosef OB (1999) Responses of auditory-cortex neurons to structural features of natural sounds. *Nature* 397:154–157.
- Phillips JA, Noppeney U, Humphreys GW, Price CJ (2002) Can segregation within the semantic system account for category-specific deficits? *Brain* 125:2067–2080.
- Rauschecker JP (1998) Cortical processing of complex sounds. *Curr Opin Neurobiol* 8:516–521.
- Reide T, Herzog H, Hammerschmidt K, Brunnberg L, Tembrock G (2001) The harmonic-to-noise ratio applied to dog barks. *J Acoust Soc Am* 110:2191–2197.
- Rizzolatti G, Fogassi L, Gallese V (2002) Motor and cognitive functions of the ventral premotor cortex. *Curr Opin Neurobiol* 12:149–154.
- Rushworth MFS, Krams M, Passingham RE (2001) The attentional role of the left parietal cortex: the distinct lateralization and localization of motor attention in the human brain. *J Cogn Neurosci* 13:698–710.
- Schroeder CE, Foxe JJ (2002) The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res Cogn Brain Res* 14:187–198.
- Schroeder CE, Lindsley RW, Specht C, Marcovici A, Smiley JF, Javitt DC (2001) Somatosensory input to auditory association cortex in the macaque monkey. *J Neurophysiol* 85:1322–1327.
- Schwartz DA, Howe CQ, Purves D (2003) The statistical structure of human speech sounds predicts musical universals. *J Neurosci* 23:7160–7168.
- Sigl JC, Chamoun NG (1994) An introduction to bispectral analysis for the electroencephalogram. *J Clin Monit* 10:392–404.
- Talairach J, Tournoux P (1988) *Coplanar stereotaxic atlas of the human brain*. New York: Thieme Medical.
- Thompson-Schill SL (2003) Neuroimaging studies of semantic memory: inferring “how” from “where.” *Neuropsychologia* 41:280–292.
- Tranel D, Damasio H, Eichhorn GR, Grabowski TJ, Ponto LLB, Hichwa RD (2003) Neural correlates of naming animals from their characteristic sounds. *Neuropsychologia* 41:847–854.
- Van Essen DC (1997) A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature* 385:313–318.
- Van Essen DC (2003) Organization of visual areas in macaque and human cerebral cortex. In: *The visual neurosciences* (Chalupa L, Werner JS, eds), pp 507–521. Cambridge, MA: MIT.
- Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH (2001) An integrated software suite for surface-based analyses of cerebral cortex. *J Am Med Inform Assoc* 8:443–459.
- Warren RM (1999) *Auditory perception*, Chap 1. Cambridge, UK: Cambridge UP.
- Wessinger CM, VanMeter J, Tian B, Van Lare J, Pekar J, Rauschecker JP (2001) Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *J Cogn Neurosci* 13:1–7.