

Interhemispheric Differences in Auditory Processing Revealed by fMRI in Awake Rhesus Monkeys

Olivier Joly¹, Franck Ramus², Daniel Pressnitzer³, Wim Vanduffel^{1,4,5} and Guy A. Orban¹

¹Laboratorium voor Neuro- en Psychofysiologie, K.U. Leuven Medical School, B-3000 Leuven, Belgium, ²Laboratoire de Sciences Cognitives et Psycholinguistique (EHESS/CNRS/DEC-ENS), Paris F-75005, France, ³Département d'Études Cognitives, École Normale Supérieure, Paris F-75005, France, ⁴Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA 02129, USA and ⁵Department of Radiology, Harvard Medical School, Boston, MA 02115, USA

Vanduffel and Orban contributed equally

Address correspondence to Dr Olivier Joly, NeuroSpin, Institut National de la Santé et de la Recherche Médicale, CEA Saclay, 91191 Gif-sur-Yvette Cedex, France. Email: olivier.joly@gmail.com.

Lesion studies in monkeys have suggested a modest left hemisphere dominance for processing species-specific vocalizations, the neural basis of which has thus far remained unclear. We used contrast agent-enhanced functional magnetic resonance imaging to map the regions of the rhesus monkey brain involved in processing conspecific vocalizations as well as human speech and emotional sounds. Control conditions included scrambled versions of all 3 stimuli and silence. Compared with silence, all stimuli activated widespread parts of the auditory cortex and subcortical auditory structures with a right hemispheric bias at the level of the auditory core. However, comparing intact with scrambled sounds revealed a leftward bias in the auditory belt and the parabelt. The left-sided dominance was stronger and more robust for human speech than for rhesus vocalizations and hence does not reflect conspecific call selectivity but rather the processing of complex spectrotemporal patterns, such as those present in human speech and in some of the rhesus monkey vocalizations. This was confirmed by regressing brain activity with a model-derived parameter indexing the prevalence of such patterns. Our results indicate that processing of vocal sounds in the lateral belt and parabelt is asymmetric in monkeys, as predicted from lesion studies.

Keywords: auditory cortex, conspecifics, functional imaging, spectrotemporal complexity, vocalizations

Introduction

Nonhuman primates, like many other animals, use vocal signals to mediate social interactions with conspecifics (Hauser 1996; Bradbury and Vehrencamp 1998). While we are beginning to understand how conspecific calls are processed in the monkey auditory system (Rauschecker and Scott 2009; Romanski and Averbeck 2009 for reviews), and how visual and auditory call signals are combined (Ghazanfar et al. 2008; Schroeder et al. 2008; Kayser and Logothetis 2009), it remains unclear whether the processing of vocalizations is left lateralized in the monkey, as is speech processing in humans. In humans, left dominance is stronger for speech understanding (Scott et al. 2000; Narain et al. 2003; Spitsyna et al. 2006; Dhanjal et al. 2008) than perception of speech (Dehaene-Lambertz et al. 2005). Initial processing of speech in auditory cortex is actually bilaterally symmetric (Poepfel et al. 2004).

A lesion study in Japanese macaques has demonstrated that the auditory cortex, especially the left cortex, is required for coo-call discrimination (Heffner HE and Heffner RS 1984, 1986). This left dominance is consistent with a right-ear advantage for

this discrimination (Petersen et al. 1978). Bilateral auditory cortical lesions also temporarily impair the monkey's ability to discriminate frequency changes, a sensory deficit that can account for the impairment of coo-call discrimination (Harrington et al. 2001). Further behavioral studies using the orienting response have yielded conflicting results in different monkey species (Hauser and Andersson 1994; Ghazanfar et al. 2001; Gil-da-Costa and Hauser 2006; Teufel et al. 2010), and this response has proven (Fischer et al. 2009) an unreliable marker for the left dominance often reported in human speech studies.

Anatomical and neurophysiological studies, focusing on the rhesus monkey, have shown that the auditory system is hierarchically organized with an auditory core (areas A1, R, RT) surrounded by 8 "belt" and at least 2 "parabelt" areas (Kaas and Hackett 2000; Hackett et al. 2001; Petkov et al. 2006). Several studies have suggested a scheme involving 2 parallel "what" and "where" auditory pathways (for review, Rauschecker and Scott 2009). It has been proposed that conspecific vocalizations are processed in the what pathway (Rauschecker and Tian 2000; Tian et al. 2001), but as far as we are aware, no asymmetric processing has been reported in this pathway.

Poremba et al. (2004) used positron emission tomography (PET) in 8 rhesus macaques (5 females and 3 males) to investigate which temporal regions underlie the asymmetry demonstrated by the lesion work of Heffner HE and Heffner RS (1986). Significantly higher metabolic activity was observed in the left than in the right temporal pole for macaque, but not human, vocalizations. In contrast to this PET study, a recent functional magnetic resonance imaging (fMRI) study (Petkov et al. 2008; data from 5 anesthetized and 2 awake male macaques) revealed a preference for macaque vocalizations in an anterior superior temporal plane region, with no asymmetry (their Supplementary Fig. S4). Similarly, Gil-da-Costa et al. (2004, 2006), using only 2 types of vocalizations, coos and screams, did not show clear-cut lateralization effects with PET imaging in 3 animals (1 male and 2 females). Thus, the imaging studies in rhesus monkeys have not revealed any consistent left dominance for the processing of vocalizations, as had been implied by the lesion results (Heffner HE and Heffner RS 1986).

To address these ambiguities, we used contrast agent-enhanced full-brain fMRI (Vanduffel et al. 2001) to investigate the processing of vocalizations in rhesus monkeys. Since the lesion literature also suggested a left lateralization for the processing of human speech (Dewson et al. 1969, 1970; Cowey and Dewson 1972), we presented a large range of monkey vocalizations, along with human speech and emotional sounds.

Importantly, we used scrambled controls preserving the long-term spectrum of the original sound, for each of these 3 sound classes. This enabled us to test the hypothesis of an early auditory cortical asymmetry and to address the additional issue of whether the asymmetry reflects low-level acoustic features of the vocalizations or aspects of their vocal nature. An acoustic measure of complexity was designed for our sound set, which was sensitive to both the spectral structure and temporal modulations of the various stimuli. This “rate/scale” metric, described in detail in Materials and Methods, was found to underlie the left hemisphere dominance consistently observed in the auditory lateral belt and parabelt.

Materials and Methods

Subjects

Three young adult rhesus monkeys (*Macaca mulatta*), 1 female (M13) and 2 males (M14, M18), 5–6 years of age, weighing between 4 and 5 kg, participated in the experiment. They were born in captivity and housed in large rooms fitted with cages for 11 pair- and group-housed monkeys. It has been shown that laboratory-housed rhesus monkeys classify, in the absence of training, species-specific calls in a manner comparable with rhesus monkeys living under more natural conditions (Gifford et al. 2003). The monkeys were experienced in viewing many types of visual stimuli but had limited exposure to experimental auditory stimuli. Before scanning sessions, they were trained daily in order to adapt them to the headphones and to increasing stimulus sound levels while performing a fixation task. The fixation task was used to equalize attention across conditions and minimize body movement during the scanning. In the fixation task, the monkey was rewarded at increasingly shorter intervals for continuing fixation within a single trial (up to 10 s). During the training, the sounds played included tones, bandpass noises, monkey calls, and human vocalizations. However, the sequences used in the scanner were different from those during the training. The details concerning head

post surgery have been previously described (Vanduffel et al. 2001; Nelissen et al. 2006). Animal care and experimental procedures met the national and European guidelines and were approved by the local ethical committee.

Experimental Design and Stimuli

The stimuli were defined by a 3×2 factorial design with 2 factors: sound class (3 levels) and scrambling (2 levels). The 3 classes of sounds (Fig. 1) that were used were human emotional (He) vocalizations, human speech (Hs), and monkey vocalizations (Mv). These 6 conditions were presented in blocks of 40 s during scanning. Within a scanning block, sequences of 2 sounds were presented within the 2.8-s intervals between 2 acquisitions. Within each block, 8 sequences were presented. Each time series or “run” included 2 blocks of the 6 conditions.

Monkey vocalizations uttered by several individuals of both sexes were drawn from the Rhesus Monkey Repertoire recorded by Marc Hauser in Cayo Santiago, Puerto Rico. Five types of social calls were selected that were described as having either positive valence (coos, girneys, and harmonic arches) or negative valence (screams and shrill barks). Calls of a given valence were concatenated into 9 positive and 10 negative sequences of 2 calls that were used in alternate runs. In the first block of a run, 8 of these sequences were used. In the second block, 7 positive and 6 negative sequences corresponded to the reversed order of the sequences presented in the first block, while the remaining sequences (1 positive and 2 negative) corresponded to the sequences not used in the first block. Sequences had a mean total duration of 2110 ± 233 ms (mean duration calls 940 ms [± 257 ms], mean interval duration 230 ms). We note here that although individual rhesus monkeys do not naturally produce sequences such as these and generally produce only a single call at a time, monkeys often collectively produce sequences of calls, for example, when moving as a troop or when anticipating food. Sequences, however, were used primarily to optimize the imaging results and to create comparable stimuli across each of the sound classes.

Human speech (Hs) stimuli were drawn from both laboratory recordings and movie soundtracks. The stimuli were uttered by several French native speakers of both sexes. In an attempt to match the

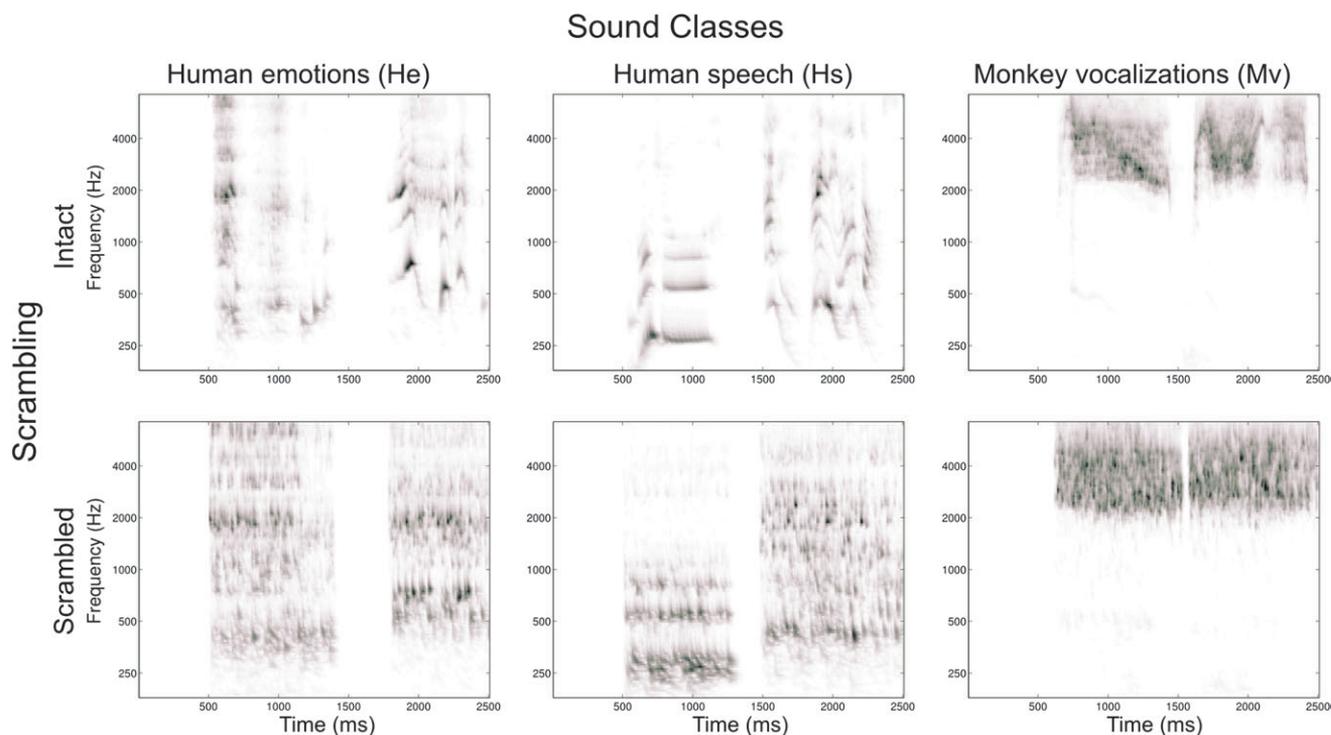


Figure 1. Auditory spectrograms. The spectrotemporal patterns (log frequency vs. time) in the output of a cochlear filtering simulation are illustrated, for a representative sound example belonging to each of the 3 classes used in the experiment: human emotional vocalizations (He), human speech (Hs), and monkey vocalizations (Mv). The spectrograms are also illustrated for the corresponding scrambled controls (bottom row).

typical brevity of the monkey calls, only single words or very short phrases (such as “hello” or “some more cake?” in French) were selected, with no particular emotional valence. They were then concatenated into 9 sequences of 2 stimuli. In the first block of a run, 8 of these sequences were presented. In the second block of a run, 7 sequences corresponded to the reversed order of the sequences presented in the first block, the eighth one corresponded to the sequence not used in the first block. Sequences had a mean total duration of 2134 ± 214 ms (mean stimulus duration 938 ms [± 149 ms], mean interval duration 258 ms). It should be noted that the monkeys had little or no prior exposure to the French language, as this is not the main language spoken in the laboratory or animal facilities. Given that monkeys do not understand the semantic content of human speech (Hs) and that monkey vocalizations (Mv) generally include a clear emotional component, we hypothesized that human emotional vocalizations might be a closer human analogue to monkey vocalizations (Belin et al. 2008). Therefore, we selected a set of human emotional (He) vocalizations, with either a positive (e.g., laughter, contentment) or a negative valence (e.g., cries, shouts), uttered by the same speakers as the speech stimuli, and that did not contain any identifiable phonetic element. Note that the definition of the valence is only from the human perspective and that we have no evidence of the possible interpretation given by the monkeys. They were again concatenated into 9 positive and 10 negative sequences of 2 stimuli. For these He stimuli, the same procedure for generating 8 sequences for each of the 2 blocks of a run was adopted as that used for the Mv stimuli. The mean total duration of human emotional sequences was 2155 ± 202 ms (mean stimulus duration 936 ms [± 266 ms], mean interval duration 283 ms). To control for different acoustic parameters, we created a “scrambled” control for each of the 3 sound classes (SHe, SHs, and SMv). Scrambled sounds were made by processing all individual (intact) stimuli through a gammatone filterbank (Patterson et al. 1995) with 64 channels. As in Patterson et al. (1995), the filterbank was chosen to mimic human frequency selectivity. The equivalent rectangular bandwidth (ERB) of each channel was thus set to $ERB = 24.7(1+4.37F)$, with F being the center frequency in kHz. This choice was motivated by the observation that macaque monkeys have a peripheral frequency selectivity that appears to be comparable with that of humans (Serafin et al. 1982; Ruggero and Temchin 2005). In each channel, the signal was windowed with overlapping Hanning windows of 25-ms duration. The windows were then shuffled randomly within a channel, with the additional constraint that a window could be displaced by no more than ± 500 ms from its original temporal position. The scrambled signals were finally obtained by putting all frequency channels back together. Scrambled speech signals have previously been used as controls for human fMRI experiments (Belin et al. 2000). In contrast to the spectral randomization of Belin et al. (2000), our method produces an exact match of spectral excitation patterns between original and scrambled signals (Fig. 2) while still making speech totally unintelligible (see Supplementary Material). Scrambled stimuli were concatenated into sequences in the same order as the stimuli of the original sequences, in order to obtain exact scrambled counterparts of the sequences in the intact conditions. Sequences of the different conditions did not significantly differ in terms of duration. The root mean square (RMS) amplitude, in arbitrary units, averaged 0.155 (standard deviation [SD] 0.008) for intact and scrambled monkey calls, 0.150 (SD 0.007) for intact and scrambled human speech, and 0.145 (SD 0.008) for intact and scrambled human emotional stimuli. A two-way analysis of variance (ANOVA) of the RMS of the 6 types of sound with scrambling and sound class as factors yielded a small but significant main effect of sound class ($F_{2,308} = 20.5$, $P < 10^{-7}$), no main effect of scrambling, and no interaction. However, the size of the effect was small, with an average-level difference of just 0.3 dB between Mv and Hs and 0.6 dB between Mv and He. These values are smaller than the just-noticeable difference for sound level in humans (Jesteadt et al. 1977).

Acoustical Analyses

All stimuli were analyzed with the model of auditory processing described in Chi et al. (2005). The model measures the spectral and temporal modulations present in sounds, using spectrotemporal filters resembling the receptive fields of A1 cells (Depireux et al. 2001). The

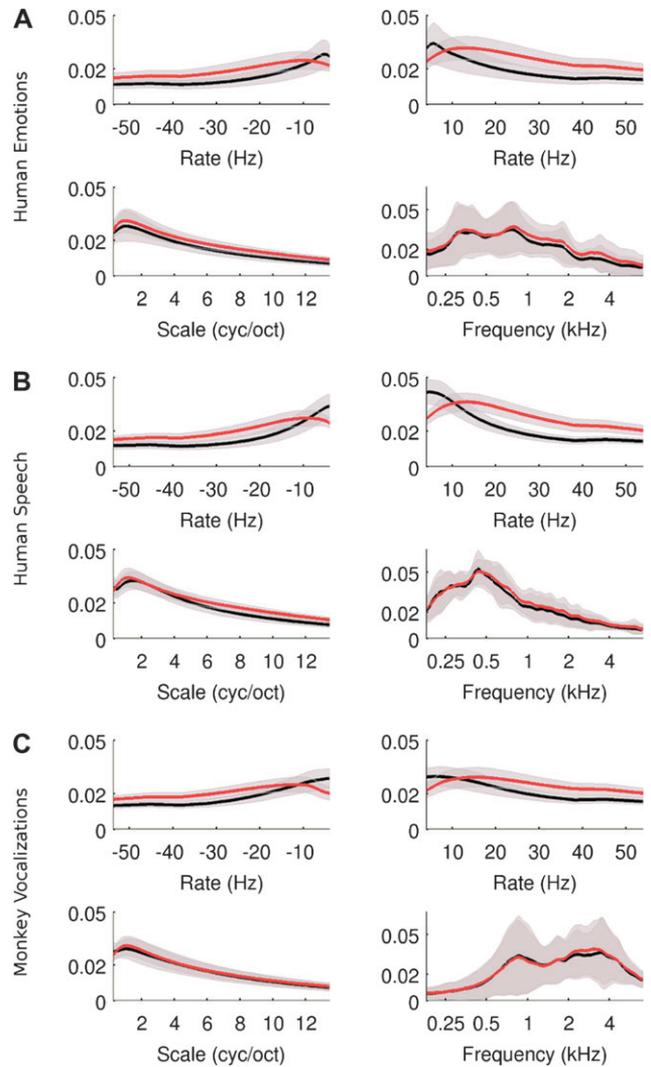


Figure 2. Auditory modeling of the stimuli. The average output of the cortical stage of the model of Chi et al. (2005) is represented for each sound class: He (A), Hs (B), and Mv (C). In all panels, the vertical axis represents mean power over the stimulus set, black and red lines represent the mean values across intact and scrambled stimuli, respectively, and the shaded area represents 1 SD around the mean. The 4D output of the model has been collapsed over time. The different panels represent averages over the other dimensions. Upper panels display “rate,” an index of the temporal modulations within the sound with high rate values corresponding to fast modulations. Negative and positive rates represent upward and downward frequency modulations, respectively. The lower left panels represent “scale,” an index of the bandwidth of spectral features, with high scales for fine spectral details. The lower right panels represent “frequency,” similar to the average excitation pattern after cochlear filtering.

initial stage of the model is an auditory spectrogram that reproduces the effects of peripheral cochlear filtering: The acoustic signal is parsed into adjacent frequency channels, corresponding to the tonotopic organization observed in all mammals (Ruggero and Temchin 2005). Examples of auditory spectrograms for the stimuli used here are given in Figure 1.

The next stage of the model, the “cortical” stage, applies spectrotemporal filters to the auditory spectrogram. These filters detect the presence of local modulations along the spectral axis (e.g., formants) or time axis (e.g., variations in amplitude) in the auditory spectrogram. The cortical stage has 4 dimensions: time, frequency, scale, and rate. Time represents stimulus time and frequency represents frequency channel, just as in the auditory spectrogram. Scale indexes the bandwidth of the spectral modulations; it is measured in cycles per

octave. Speech, with its sharp peaks and troughs of energy in different frequency channels, typically has a high scale value, whereas noise has a low scale. Rate indexes temporal envelope modulations in the auditory spectrogram and is measured in Hertz. Fast variations in amplitude produce high rates. Moreover, for joint spectrotemporal modulations such as variations in frequency, rates can be positive or negative: A downward frequency modulation produces a positive rate, whereas an upward modulation has a negative rate. A pure amplitude modulation within one channel has equal positive and negative rates. More details about the model, including examples of analyses for various kinds of sounds, are available in Chi et al. (2005).

The full model has a 4D output, which is difficult to use for regressing brain-imaging data. We therefore devised a novel statistic to summarize the full model output for a given sound. The aim was to identify sounds that, like speech, combine fine spectral details with slow temporal modulations. Stimuli were first passed through the full cortical model (all parameters of the analysis are given in Supplementary Table 1). The output of the spectrotemporal filters was averaged across time, frequency, and upward and downward modulations, and the center of mass of this representation was computed to estimate the dominant scale and rate present over the whole time course of each sound file. In some analyses, the dominant rate was used as an index. For other analyses, a rate/scale statistic was computed by taking the ratio between the dominant rates and scales. The reasoning was as follows: Low values of the rate/scale index should be obtained if the sound contains slow temporal modulations (low rate) combined with fine spectral structure (high scale), as is the case with speech (see also Acoustic Characterization of the Stimulus Set). The rate/scale ratio can thus be thought as a quantitative estimate of spectrotemporal complexity, a factor that appears important for the functional organization of human cortical processing of sounds (Samson et al. 2011).

Acoustical Characterization of the Stimulus Set

The cortical model's output for the different classes of stimuli used in the experiment is illustrated in Figure 2 for intact (black lines) or scrambled (red lines) sounds. Frequency and scale distributions are matched in intact and scrambled sounds, as expected, with perhaps a tendency to lower scales after scrambling. There is, however, a mismatch for the rate parameter. This was also expected from the scrambling algorithm: Shuffling short time windows disrupted any slow amplitude modulations and introduced higher rates. Rate was significantly higher (paired *t*-test, all $P < 10^{-13}$) for scrambled compared with intact sounds for all 3 sound classes. A two-way ANOVA with scrambling and sound class revealed extremely significant main effects of sound class ($F_{2,308} = 53$, $P < 10^{-15}$) and of scrambling ($F_{1,154} = 608.4$, $P < 10^{-15}$) but no interaction.

We verified that the rate/scale index could identify the speech-like sounds in our stimulus set. The rate/scale index is plotted in Figure 3 for the different stimulus classes and for intact and scrambled sounds. Sound class had an effect on the rate/scale index. Hs generally produced low rate/scale because of the presence of fine spectral details and slow modulations. The rate/scale indices for He were generally higher than those for Hs because of the slower modulations of speech. Mv yielded a bimodal distribution. Some calls had low rate/scale indices, comparable with human vocalizations, while other vocalizations had the highest rate/scale indices of the entire stimulus set. This reflects the heterogeneity of monkey vocalizations audible to the casual observer: The low rate/scale indices correspond to calls that have a distinct vocal quality (typically, coos, most girneys, and some screams), while the high rate/scale indices correspond to calls that have a more noise-like quality (typically, shrill barks, a few girneys, and some screams). Calls have previously been classified into 3 categories (tonal, harmonic, and noisy) on an acoustic basis (Rauschecker 1998); here, we use rate/scale as a quantitative metric to describe the acoustic features of calls. This characterization was confirmed by the analysis of a larger set of 66 calls comprising all 5 types, from which the experimental stimuli were selected: The median rate/scale was 8.1 oct/s for coos, 8.3 oct/s for girneys, 9.9 oct/s for harmonic arches, 10.7 oct/s for screams, and 12.9 oct/s for shrill barks (see Supplementary Fig. S1). This dissociation between coos and shrill barks is very reminiscent of the dendrograms based on acoustic features computed by Averbeck and Romanski (2006).

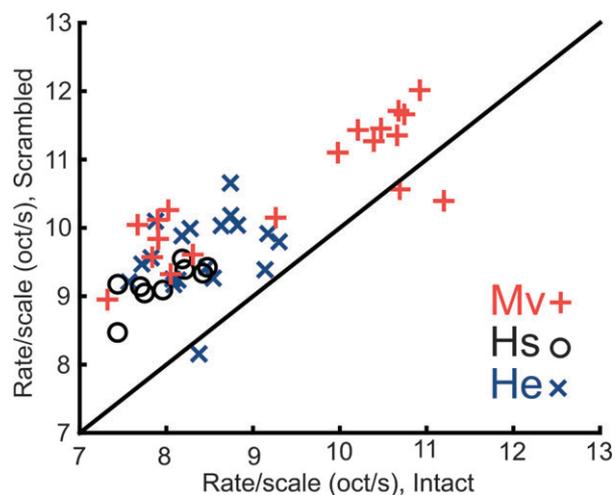


Figure 3. rate/scale index. The scatter plots for the 3 sound classes: He (blue multiplication symbols), Hs (black circles), and Mv (red crosses) represent the paired rate/scale indices (see Materials and Methods) for the scrambled and the intact version of all the sound sequences (19 for He and Mv and 9 in panel Hs). Note that sequences based on the reversed order (second scanning block) of original sequences yield the same data points.

Scrambling also had an effect on the rate/scale index. As can be seen in Figure 1, scrambling disrupted slow modulations, if they were present in the intact stimulus, and tended to smear spectral features. This was reflected in the higher rate/scale indices for scrambled than for natural stimuli. A paired *t*-test confirmed significant rate/scale differences between intact and scrambled stimuli for Hs ($t = 13.9$, $P < 10^{-6}$) and He ($t = 8.5$, $P < 10^{-7}$), as well as for Mv ($t = 6.6$, $P < 10^{-5}$). A two-way ANOVA revealed significant main effects of sound class ($F_{2,308} = 34.8$, $P < 10^{-12}$) and scrambling ($F_{1,154} = 83.3$, $P < 10^{-15}$). The interaction was not significant, indicating that scrambling influenced rate/scale similarly for the 3 classes. The variance of the differences between intact and scrambled sounds, however, was significantly larger for Mv than for Hs ($F = 8.7$, $P < 0.05$). Thus, the nature of the intact stimuli influenced the strength of the effect of scrambling.

To summarize, scrambling largely destroyed the complex spectrotemporal patterns that characterize human vocalizations and some of the monkey vocalizations. The rate/scale index captures the disruptive effect of scrambling on those sounds. Importantly, the rate/scale parameter is also sensitive to the differences between intact stimuli: Monkey vocalizations span a wide range of rate/scale, almost as broad as the effect of scrambling, depending on whether or not they are speech-like. We will therefore use the rate/scale index as a parameter of interest to localize brain areas implicated in complex spectrotemporal processing.

Magnetic Resonance Imaging Acquisition and Sound Presentation

During scanning, the monkeys sat in a sphinx position within the magnet, facing a screen onto which a red fixation point was projected (Barco LCD projector). The position of one eye was monitored at 120 Hz using a pupil-corneal reflection tracking system (Iscan, Inc., MA, USA). Monkeys received a juice reward for maintaining fixation within a small window centered on the fixation target. Before each scanning session, monocrySTALLINE iron oxide nanoparticle contrast agent (MION; Sinerem) was injected into the saphenous or femoral vein (4–10 mg/kg) to increase the contrast-noise ratio and improve the localization of the signal (Vanduffel et al. 2001; Leite et al. 2002). Monkeys were scanned in a horizontal 1.5T scanner (Sonata; Siemens Medical Solutions, Erlangen, Germany) using a receive-only surface coil positioned over the head.

In a block design, each functional time series defined a “Clustered Volume Acquisition” scheme (Kovacs et al. 2006) and consisted of gradient-echo echo-planar whole-brain images (EPs): repetition time (TR) = 5 s; acquisition time = 2.2 s; echo time = 27 ms; slices thickness = 2 mm; field of view = 128 mm; matrix size 64 × 64 yielding a resolution of

2 × 2 × 2 mm. Intact and scrambled monkey vocalization (Mv), human speech (Hs), and human emotional (He) stimuli were presented to both ears simultaneously, for about 2 s, in the silent gap (2.8 s) between the acquisitions of 2 functional volumes. Each time series included 2 presentations of each condition (6 sound conditions and a silent baseline) in blocks of 8 TR (40 s), the order of which was randomized across time series. A time series therefore lasted 560 s (i.e., 9 min 20 s). A total of 10 752 (112 × 96) volumes were acquired across all scanning sessions. Based on the quality of fixation maintained by the monkey, a subtotal of 9408 (112 × 84) volumes (EPI) entered the group analysis. Single-subject analysis included 3136 (112 × 28) volumes per subject.

Sound sequences, saved as “wav” files (sampling frequency = 22 050 Hz), were played with custom software and delivered using magnetic resonance (MR)-compatible headphones (Baumgart et al. 1998) integrated into ear mufflers designed for passive gradient noise dampening and customized for monkeys (MR Confon GmbH, Magdeburg, Germany). These headphones minimize the distortion of sounds delivered at the ear. Sound intensity measurements were made with a microphone and a sound level meter (Bruel & Kjaer GmbH, Bremen, Germany). Stimuli were presented at ~80 dB sound pressure level (SPL). The scanner noise was measured to reach up to 93 dB SPL, but, given the -20-dB attenuation by the headphone cups, the scanner noise reaching the monkey’s ears is estimated at less than 73 dB SPL.

Data Analysis

Time series were analyzed using adapted SPM5 software (<http://www.fil.ion.ucl.ac.uk/spm/>). Spatial preprocessing consisted of realignment and rigid coregistration with a template anatomy (M12; 0.35 × 0.35 × 0.35 mm voxels) in stereotaxic space. To compensate for echo-planar image distortion and interindividual anatomical differences, functional images were warped to the template using a nonrigid matching technique (BrainMatcher software; INRIA). The images were resampled to 1 mm isotropic and finally smoothed with an isotropic Gaussian kernel (full-width at half-maximum = 1.5 mm).

Fixed-effect group analysis was performed with an equal number of 28 runs per monkey. The same 28 runs were used for single-subject analysis. In one SPM analysis, the 6 sound conditions and a silent baseline entered the General Linear Model (GLM), and the realignment parameters were included as covariates of no interest. In another SPM analysis, we targeted the regions that correlate with the rate/scale ratio derived from the output of the cortical stage. The associated regressor was defined by the mean rate/scale value of each stimulus pair, convolved with the (MION) hemodynamic response function and subsampled to the TR. In this SPM analysis, the GLM included 1) the alternation between sound stimuli and silence, 2) the realignment parameters as regressors of no interest, and 3) the rate/scale associated regressor to target regions of interest. We addressed the reliability of the analysis of this regression by splitting the data set into 2. The 2 data sets included data from the same 3 individuals but on different scanning days. A final group analysis was performed using rate as regressor rather than rate/scale.

We further assessed the robustness of the experimental results by complementing the group analyses with single-subject analyses. SPM maps were either projected onto a flattened cortical surface using caret software (brainmap.wustl.edu) or overlaid onto the high-resolution anatomical magnetic resonance imaging (MRI) of our template M12 using Anatomist (<http://brainvisa.info>, last accessed february 6, 2011). Thresholds for the group and single-subject analyses, including the regression with rate/scale were set at $P < 0.05$, familywise error (FWE) corrected for multiple comparisons (using Random Field Theory), unless specified otherwise.

To facilitate further cross-study comparison, the M12 (=M1 from Ekstrom et al. 2008) template anatomy (after skull stripping) was registered to the population-average MRI-based template for rhesus macaque, later referred to as 112RM-SL (McLaren et al. 2009), which is furthermore aligned to the MRI volume from a histological atlas (Saleem and Logothetis 2006). This registration was performed using the nonrigid symmetric diffeomorphism approach (SyN) implemented in the ANTS (version 0.5) package (Avants et al. 2008). The choice of this approach was instigated by a recent evaluation of 14 nonlinear deformation algorithms revealing that SyN combines flexibility with high accuracy (Klein et al. 2009). Indeed, the difference between the 2

templates (M12 and 112RM-SL) is more than a simple translation and can probably be explained by age differences between M12 and the subjects contributing to 112RM-SL space. Hence, in this report, xyz stereotaxic coordinates are given in 112RM-SL space unless specified otherwise. The registration to 112RM-SL space allowed us to indicate the borders of atlas-defined regions such as A1, the border between the caudal belt and area Tpt, and the border of the anterior core and belt with area RTp.

Effect of Hemisphere and Lateralization Indices

Cerebral hemispheric specializations were assessed by means of 2 complementary approaches. First, we statistically tested for significant left-right asymmetries by entering the original EPIs and their flipped versions into an SPM analysis. In this analysis, the effect of “hemisphere,” comparing left and right hemispheres, can be tested for any contrast. This interaction tests, for each voxel, the significance of the difference between the contrast in that voxel and in the corresponding voxel in the opposite hemisphere. The threshold is set at $P < 0.001$ (t value = 3.09), uncorrected for multiple comparisons. It is worth mentioning that in calculating interactions, variances are added and that this threshold is therefore rather stringent (Georgieva et al. 2009).

To further investigate hemispheric specialization, fMRI-derived lateralization indices (LIs) were calculated by means of the LI toolbox for SPM (Wilke and Lidzba 2007) using the following options: ±5 mm midsagittal exclusive mask, clustering with a minimum cluster size of 5 voxels, and default bootstrapping parameters (min/max sample size: 5/10 000 and bootstrapping sample size set to 25% of input size). Since the activation patterns were relatively large and confined to the lateral parts of the hemispheres, the exact choice of the mask and clustering parameters was not critical. LIs were calculated on the basis of t -contrasts, integrating the sum of voxels in each hemisphere considering only above-threshold values. Hence, LIs were derived from the following equation: $LI = (\text{right} - \text{left}) / (\text{right} + \text{left})$, which leads to negative values for predominantly left hemisphere activation. The LI curve plots the bootstrapped LI values as a function of the statistical threshold (mean value in white, supplemented by minimum and maximum bootstrapped LI in color). An overall weighted bootstrapped LI can be calculated for each contrast and for each individual.

Analysis of Eye Position Recordings

The position of one eye was monitored at 120 Hz during scanning. Monkeys received a juice reward for maintaining fixation within a small window centered on the fixation target. Percent fixation was computed as the ratio between the time spent within the 2° fixation window and the total duration of the block (40 s). Horizontal and vertical SDs of the traces were calculated for the time spent within the fixation window (Joly et al. 2009). Saccades were detected as portions of the traces that were associated with 1) instantaneous speed that was 3 SD or more above the mean and 2) eye position outside the fixation window for more than 60 ms. Significant ($P < 0.05$) differences between conditions for these parameters were assessed by a one-way ANOVA.

Results

Fixation Behavior

On average, monkeys held their gaze in the fixation window for more than 90% of the time in each run included in the analysis, and the percent of fixation was not significantly different across conditions: 96.55% (M13, $P = 0.798$), 96.53% (M14, $P = 0.706$), and 95.17% (M18, $P = 0.817$). No differences were observed in the number of saccades or in horizontal and vertical SDs across conditions, for any of the individuals.

Auditory Activation

To identify the brain regions involved in auditory processing, we computed the main auditory activation using the contrast (all sounds - silence) for the group of 3 monkeys. Figure 4A shows the resulting SPM t -maps with the significant voxels projected

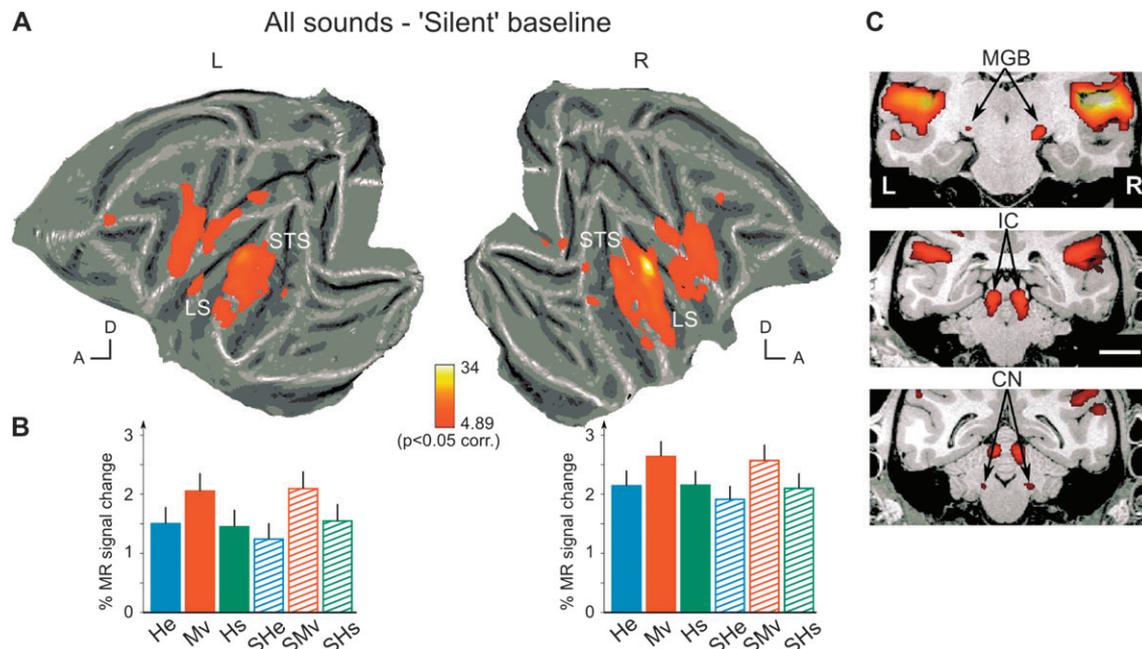


Figure 4. Cortical and subcortical main auditory activation in the group analysis (3 monkeys). (A) Projection of the SPM t -map ($P < 0.05$ FWE corrected, fixed effect) for the contrast “all sounds” minus “silence” onto the flattened left and right hemispheres (M12). The dark-to-white gray scale of the brain map corresponds to convexity of the cortical surface: black corresponding to the fundus of sulci, white to the crown of cortical gyri, and gray to the banks of sulci. The red color code (color bar in the middle) indicates t -scores; D: dorsal, A: anterior; (B) Activity profiles representing the percentages (mean \pm standard error of the mean across runs) of MR signal change for each condition compared with silent baseline at the global maximum of each hemisphere. (C) Overlay of the t -score map of the main auditory activation onto different coronal sections at the level of the MGB, IC, and CN; bar scale: 1 cm.

onto the flattened cortical surface of the left and right hemispheres. The auditory activation was bilateral and showed a global maximum in the right hemisphere at coordinates [18, 7, 20] in 112RM-SL space, most likely located within the primary auditory cortex. Comparison with the tonotopic regions of the atlas (Fig. 5B) suggests that the local maximum is located in A1. The percent of signal change (PSC) relative to the silent baseline is plotted as a function of stimulus condition for the global maximum of each hemisphere in Figure 4B. The average signal change extracted from the signal at voxel with the maximum of activity for each hemisphere was about 35% stronger in the right hemisphere compared with the left. The relative activity across conditions was similar in both hemispheres, showing the strongest signal changes for monkey vocalizations (Mv and SMv). Yet the profiles did not simply reflect the relative intensity of the stimuli (see Materials and Methods).

The auditory stimuli activated most of the lower bank of the lateral sulcus (LS) and extended mainly into the superior temporal gyrus (STG), roughly up to the anterior border of the rostral parabelt (RPB; Hackett et al. 1998). The activation within the upper bank of the superior temporal sulcus (STS) was often indistinguishable from strong activation in the LS, spilling over into the banks of the STS. Bilateral activations were observed more dorsally in the motor and somatosensory cortices (ventral parts corresponding to the head), in the anterior inferior parietal lobule (IPL), and in the anterodorsal insular cortex. Unilateral activations were observed in the left lateral bank of the intraparietal sulcus (IPS) and in the right occipital visual cortex.

At the level of subcortical structures (Fig. 4C), overlaying the SPM t -map with the anatomical MRI template revealed significant auditory activation of the medial geniculate bodies (MGBs) in

both the left hemisphere (coordinates -9, 7, 11; $t = 4.97$, #voxels in cluster = 1) and right hemisphere (coordinates 9, 8, 10; $t = 6.21$, #voxels = 20). There was also a bilaterally significant auditory activation in the inferior colliculi (ICs) at [-4, 0, 9] ($t = 14.8$, #voxels = 105) and at [4, 0, 9] ($t = 18.5$, #voxels = 139) and in the cochlear nuclei (CN) at [-6, -2, -1] ($t = 5.58$, #voxels = 12) and at [7, -3, 0] ($t = 5.74$, #voxels = 10).

Effect of Scrambling

To target brain regions that preferentially respond to intact vocalizations, we computed the main effect of scrambling using the contrast (“intact sounds” - “scrambled sounds”) within the main auditory activation (using an inclusive mask of “all sounds” vs. “silence” at $P < 0.05$ uncorrected). Figure 5A,B shows the voxels reaching significance, projected onto the cortical flat maps. To appreciate the spatial specificity of this activation for intact sounds compared with the main auditory activation, we overlaid the significant voxels from the 2 SPMs onto the same flat maps, showing the scrambling effect in red-yellow voxels and the main auditory activation in dark green-white voxels. The activation of the scrambling effect was restricted to the lower bank of the LS, extending into the STG. The local maximum for the intact versus scrambled vocalizations was observed within the LS, albeit more laterally than for the main auditory activation. An activation site was also observed in the left orbitofrontal cortex at the level of the lateral orbital sulcus (Fig. 5A).

Figure 5B represents an enlarged view of a portion of the flat map within the white lines on Figure 5A. These detailed flat maps include the LS, the STG, and the STS. To assist in the localization of the activations, we projected atlas-defined borders onto these maps. From posterior to anterior, we show the border between area Tpt and the caudal belt (1), the posterior

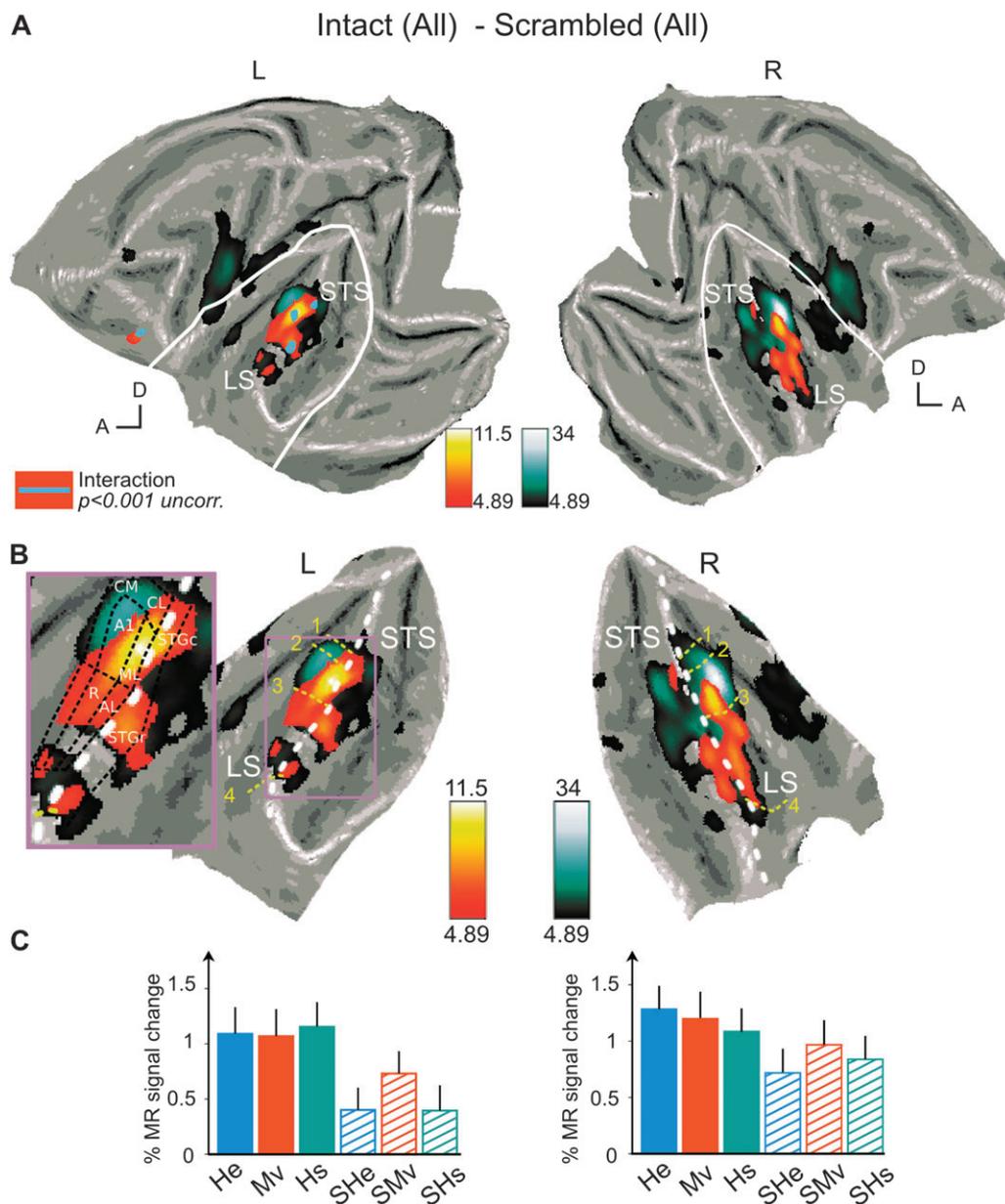


Figure 5. Main effect of scrambling and interaction with sound class (group analysis, 3 monkeys). (A) SPM t -maps (in red voxels) for the contrast “intact” minus “scrambled” sounds ($P < 0.05$ FWE corrected, fixed effect, and inclusively masked with the main auditory activation at $P < 0.05$ uncorrected) projected onto the cortical surface of both hemispheres. The main auditory activation (as in Fig. 4A) is shown in green. Color bars in the middle indicate t -scores for both contrasts. The light-blue outlines represent the significant interaction between sound class and scrambling defined as $([Hs - SHs] - [Mv - SMv])$, at $P < 0.001$ uncorrected, using the main effect of scrambling as inclusive mask. White lines represent the cuts that define the portion of the cortical surface shown in panel (B). (B) Flat maps of a subregion of the cortical surface, limited to the LS and STS, with the same functional maps as in panel (A). White dashed lines represent the lips of the LS. Yellow dashed lines represent atlas-defined borders from posterior to anterior: the border between area Tpt and caudal belt (1), the posterior (2) and anterior (3) borders of area A1, and the anterior border of auditory core and belt (4). The inset shows the right flat map in relation to the posterior tonotopic (core and belt) regions plus STG. (C) Activity profiles representing the percent (mean + standard error of the mean across runs) signal change for each condition compared with silent baseline at the global maximum of the scrambling effect in the left hemisphere and the symmetric voxel in the right hemisphere ($\pm 24, 8, 20$).

border of area A1 (2), the anterior border of area A1 (3), and the anterior border of auditory belt and the core, where they meet Ts2 (4). In addition, in the inset, the outlines of the middle and posterior part of the auditory core, belt, and STG are indicated. The main effect of scrambling reached its maximum in the left hemisphere between borders (2) and (3), a location corresponding (see inset in Fig. 5B) to the field ML of the lateral belt. The main activation extended anteriorly into AL and neighboring parabelt. Significant activation was also found in a more RPB

region and in the anterior part of the LS near the border between the auditory core and the medial belt at the level of the border between area R and RT. Interestingly, AL and ML project to the orbitofrontal cortex (Romanski et al. 1999) in the lateral orbital sulcus region.

The global maximum for the main effect of scrambling was located at $[-24, 8, 20]$. The PSC for each condition relative to the silent baseline is plotted in Figure 5C for the global maximum of the left hemisphere and its symmetric voxel in the right

hemisphere at [+24, 8, 20]. Although the overall activation was higher in the right hemisphere (paired t -test, $t = 2.19$, $P < 0.03$), the differential activity between intact and scrambled sounds was significantly larger (paired t -test, $t = 2.79$, $P < 0.005$) in the left hemisphere, as expected from the SPM analysis. Figure 5C also suggests that the scrambling effect is different for the various types of sound, as predicted by the acoustical analyses (Figs 2 and 3). In particular, the scrambling effect should be stronger for Hs than Mv in regions engaged in processing speech-like spectrotemporal patterns. Therefore, we computed the interaction between scrambling and sound class, specifically the contrast $([Hs - SHs] - [Mv - SMv])$. This interaction reached significance ($P < 0.001$ uncorrected) only in the left hemisphere, and specifically in the LS, the STG, and the lateral orbitofrontal cortex (light-blue outlines in left hemisphere of Fig. 5A). The opposite contrast $([Mv - SMv] - [Hs - SHs])$ yielded no significant voxels.

Lateralization—Effect of Hemisphere and LI Curves

The main auditory contrast as well as the scrambling effect yielded a bilateral activation pattern but with opposite asymmetries (Figs 4 and 5). The right-sided preference for the contrast all sounds versus silence (Fig. 4A) and the leftward preference for intact versus scrambled sounds (Fig. 5) were further assessed by 2 complementary approaches: 1) by testing the effect of hemisphere using statistical parametric mapping and 2) by calculating the LI.

In the first analysis, we computed the interaction of hemisphere with both the main auditory contrast and the effect of scrambling (Fig. 6A). The effect of hemisphere for the main auditory contrast (dark green voxels in Fig. 6A) reached significance only in the right hemisphere. The local maximum was located in the posterior part of the LS, near the global maximum for the contrast all sounds versus silence. The effect

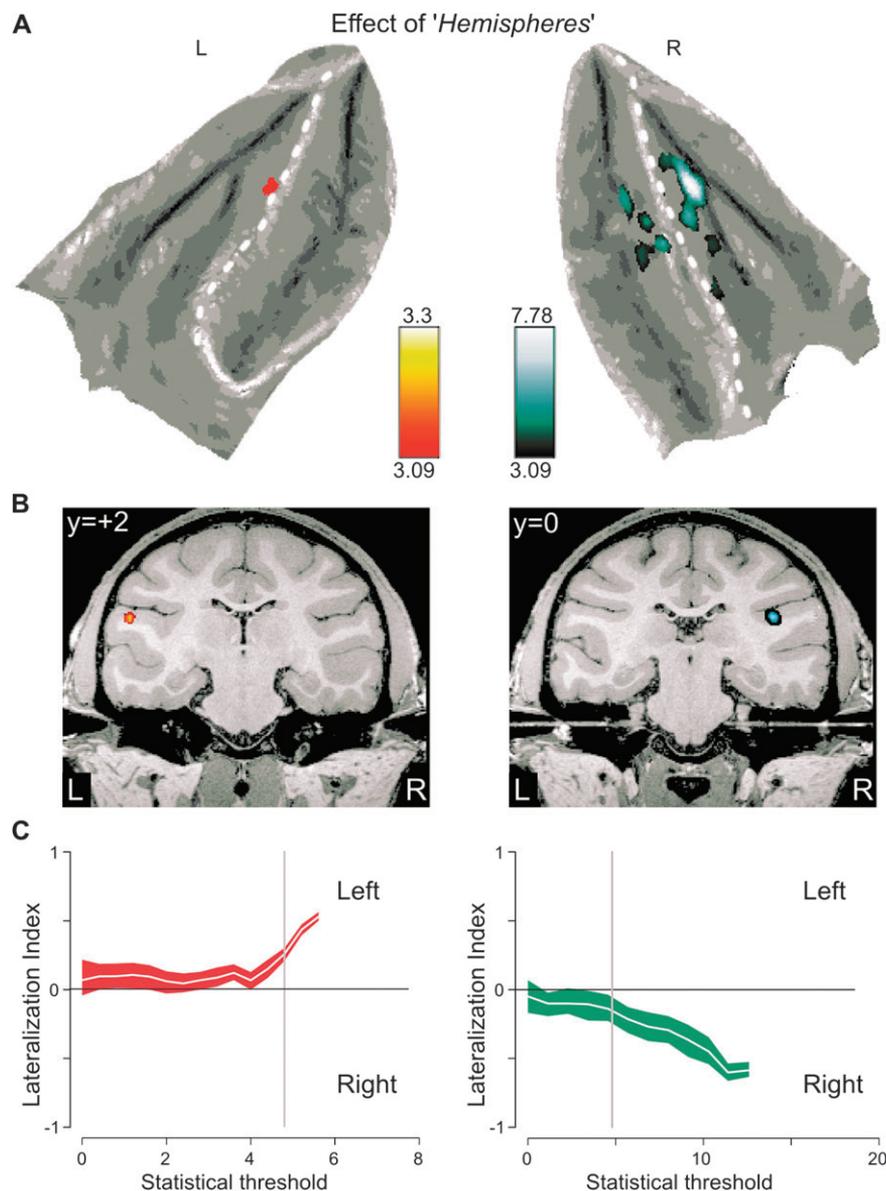


Figure 6. Lateralization of activation patterns (group analysis, 3 monkeys). (A) SPM t -maps ($P < 0.001$ uncorrected) of the interaction of the factor hemisphere with both the main auditory activation (in dark green) and the main effect of scrambling (in red) after projection onto left and right cortical surfaces containing LS and STS. (B) MRI coronal section overlaid with the effect of hemisphere onto the main effect of scrambling (left, maximum at $-24, 2, 23$ in M12 space) and the main auditory activation (right, maximum at $17, 0, 23$ in M12 space). (C) LI curves, plotting the index as a function of the statistical threshold (t -score) applied to the SPM t -maps of both the main effect of scrambling (left, weighted LI = $+0.21$) and the main auditory activation (right, weighted LI = -0.38).

of hemisphere for the main effect of scrambling (red voxels in Fig. 6A) was found only in the left hemisphere, near the lip between the LS and STG. To better characterize the localization, we overlaid both SPM *t*-maps onto the anatomical MRI of the template (Fig. 6B). The coronal sections shown at $y = 0$ and $y = +2$ (M12 anatomical space) correspond to $y + 5$ and $y + 7$ in 112RM-SL space (see Materials and Methods). The *t*-maps confirm the differences in the localizations of the scrambling effect and main auditory activation asymmetries within the lower bank of the left and right LS.

For both contrasts, we also plotted the LI curves, which represent the LI ($L - R/L + R$) as a function of the statistical threshold. Not surprisingly, this analysis revealed a right-biased LI curve (green curve in Fig. 6C), associated with a negative mean LI (-0.38) that quantifies the degree of right hemispheric preference for the main auditory activation in this data set. Conversely, a left-biased LI curve was observed for the main effect of scrambling (red curve in Fig. 6C), associated with a positive mean LI (+0.21). While the first analysis is a straightforward voxel-based test, the LI curve analysis ensures that the lateralization effect observed with the former method is not caused by slightly asymmetric activation extents in the 2 hemispheres. A very similar pattern in each hemisphere at a slightly asymmetric location could yield significant voxels in the first voxel-based analysis. However, the slightly asymmetric activation locations in the 2 hemispheres would be insufficient to generate a monotonic LI curve (based on the total sum of above-threshold voxels) or to give rise to a nonzero LI-weighted value. Together, these methods demonstrate opposite asymmetries for the main auditory activation and for the main effect of scrambling. The former right-sided effect is localized in area A1 and the latter left-sided effect in the region ML of the lateral belt.

Sound Class

To complement the information provided by maps of the main auditory activation and the main effect of scrambling, we computed the simple effect of intact sound – silent baseline (in green) and the effect of scrambling (in red) for each sound class separately (Fig. 7), using the same color conventions as in Figures 5 and 6. For each sound class, the contrast “intact sound” – “silent” baseline (in green, $P < 0.05$ corrected) reached its maximum in the right hemisphere, as did the main auditory effect illustrated in Figures 4 and 6. The effect of scrambling (in red, $P < 0.05$ corrected) reached its maximum in the left hemisphere at the level of the lateral belt for each sound class. While the scrambling effect reached significance for Hs and He (yellow and red voxels in Fig. 7A,B), the Mv class (Fig. 7C) showed the weakest effect of scrambling, and it is therefore represented by a red dashed outline corresponding to $P < 0.001$ uncorrected for multiple comparisons. Figure 7 also indicates that the right asymmetry for sound versus silence is more robust than the left asymmetry for scrambling. The former effect was equally strong for the 3 sound classes. The latter effect, on the other hand, depended on the type of sound and was much stronger for human speech than for the 2 other classes, in agreement with the predictions from the acoustical analyses (Figs 2 and 3).

In Figure 7C, a yellow dashed line indicates the anterior border of the auditory core and belt as in Figure 5B, which corresponds to $y = +23$ in 112RM-SL space ($y = +18$ in M12). This line indicates the probable posterior border of area Ts2 used as landmark in Petkov et al. (2008). Very little auditory

activation by monkey calls (green in Fig. 7C) was found anterior to this line, especially in the left hemisphere, and no significant effect of scrambling for monkey vocalizations (Mv – SMv) was observed, not even at $P < 0.001$ uncorrected (red dashed lines in Fig. 7C), in this very anterior region.

Sensitivity to the Rate/Scale Index

In order to directly visualize the regions involved in the processing of the spectrotemporal patterns characteristic of speech sounds, we used the rate/scale index described in Materials and Methods as a regressor and identified the regions in which activity increased with decreasing rate/scale (keeping in mind that low values of rate/scale correspond to speech-like sounds in our stimulus set). This approach yields a strongly left-lateralized activation pattern, with 3 sites in the left hemisphere (Fig. 8A). The most posterior of the 3 left regions was located near the ML-AL transition in the auditory belt, in the vicinity of the local maximum of the global scrambling effect (Fig. 5B). The 2 other left STG regions belonged to the RPB. The more anterior of these STG regions was the weakest site and occurred only in the left hemisphere, while the middle STG region was activated to some extent in both hemispheres (Table 1) and corresponded to the site in the STG activated by the scrambling effect for speech (Fig. 7B). This middle site was also left lateralized (Table 1) and was close to one of the local maxima for the interaction between sound class and scrambling (Fig. 5A).

The activity profile of the middle STG region (Fig. 8C) confirms that activity is lower for scrambled than for intact stimuli. The inverse relationship observed between activity in this region and the rate/scale index, which is higher here for scrambled than for intact sounds (Fig. 3), indicates that activity increases for more speech-like spectrotemporal patterns. The LI curve (Fig. 8D) confirms the strong left lateralization of rate/scale processing, with the weighted LI reaching +0.45, the highest value observed in the present study. Indeed, the weighted LI of the scrambling effect reached only +0.21.

Patterns similar to those observed in the full data set were also seen in the 2 split data sets (compare Fig. 8A and Fig. 9A,B). The maps are shown at a lower statistical threshold ($P < 0.001$ uncorrected for multiple comparison), however, because of the reduced sensitivity of half data sets. In both split data sets, the activations were stronger in the left hemisphere. Indeed, the weighted LI of the 2 independent data sets were both positive (+0.25 and +0.20) showing a consistent leftward lateralization of the rate/scale analyses. It is worth noting that the middle STG activation (arrow) was consistently present and left lateralized in both data sets.

Comparison of Figures 8A and 5B reveal that the pattern of regions correlating negatively with rate/scale is clearly different from that of the scrambling effect. Indeed, the rate/scale parameter not only depends on the scrambling effect but also reflects the spectrotemporal differences among intact stimuli, including those between the different vocalizations (see above). The activation pattern for rate/scale is also different to some degree from that obtained with rate as regressor (Fig. 8B). The latter pattern included the same 3 regions as the rate/scale activation pattern (Table 1), but by far the strongest activation was located in a more posterior region, close to the local maximum of the global scrambling effect. Furthermore, the rate pattern also extended medially toward the auditory core (Fig. 8B). Indeed, although rate depended both on sound class and scrambling, the latter

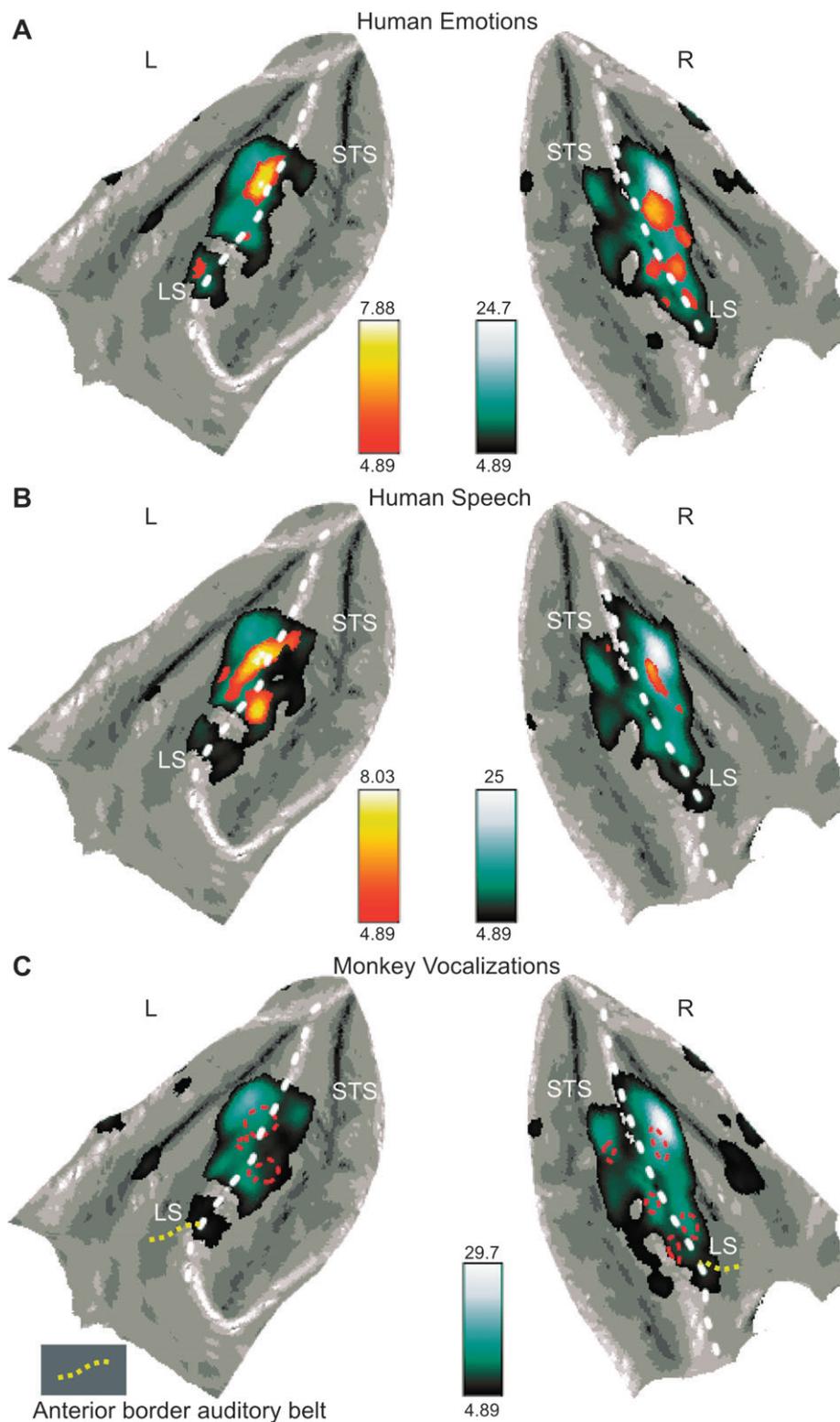


Figure 7. Auditory activation and scrambling effect for different sound classes (group data, 3 monkeys). (A–C) SPM t -maps of auditory activation (intact – silence, dark green color, $P < 0.05$ FWE corrected) and effect of scrambling (intact – scramble, red color, $P < 0.05$ FWE corrected) for human emotion (A), human speech (B) and monkey vocalizations (C). In (C), the effect of scrambling for monkey vocalizations (Mv – SMv) is shown at a lower threshold ($P < 0.001$ uncorrected) by the dashed red outlines. The yellow dashed line indicates the atlas-defined anterior border of the auditory belt for comparison with Petkov et al. (2008).

dominated. Given its more restricted activation pattern, the rate/scale index captures the higher-order auditory processing of complex spectrotemporal patterns, such as speech, better than a rate index.

Comparison across Individuals

To evaluate the consistency of the main auditory activation, the main effect of scrambling, and their respective lateralizations, we computed and displayed the SPM t -maps of each of these contrasts

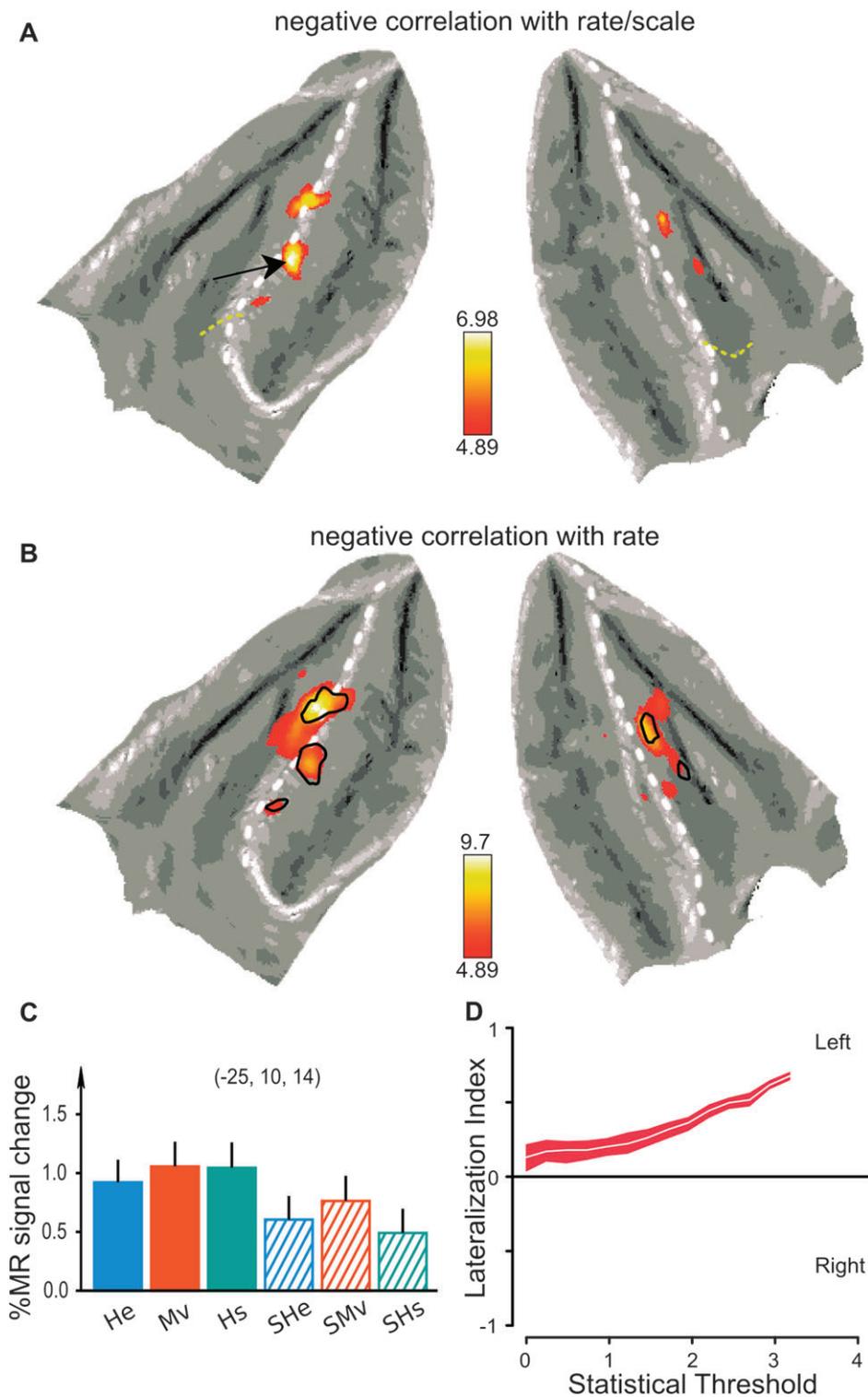


Figure 8. Effect of the rate/scale index (group analysis, 3 monkeys). (A) SPM t -map ($P < 0.05$ FWE corrected, fixed effect) defined by regressing the MR activity with rate/scale. Speech-like sounds have a low rate/scale index, hence the negative correlation between MR activity and rate/scale. (B) SPM t -map ($P < 0.05$ FWE corrected, fixed effect) defined by regressing the MR activity onto the rate of the stimuli. (C) Activity profile of the left middle STG site (arrow in A), (D) LI curve of the SPM t -map shown in (A). The yellow dashed line in (A) indicates the atlas-defined anterior border of the auditory belt for comparison with Petkov et al. (2008). The black outlines in (B) represent the rate/scale activation pattern from panel (A).

separately for the 3 individuals (see Supplementary Fig. S2). Main effects of scrambling were observed in M14 and M13 (see Supplementary Fig. S2A,C). The left lateralization of the scrambling effect was observed in the same 2 monkeys, a finding confirmed by the LI (Table 2). While in M14 and M13, the LI values were positive

for each of the single scrambling effects; this value was negative in M18. It is noteworthy that this animal (M18), which failed to show the left asymmetry for the scrambling effect, was also the one in which the right-sided asymmetry of the main auditory activation was far stronger than in the other 2 animals.

Table 1

Hemisphere, coordinates (x, y, z), group t -score, and number of voxels reaching $P < 0.05$ corrected, as well as individual t -scores for the 5 sites involved in processing spectrotemporal complexity (Fig. 8)

Hemisphere	Coordinates	Group t -score	Number of voxels	t -Score M14	t -Score M18	t -Score M13
Left	-25, 10, 14	7.39	88	6.48	4.19	5.62
Left	-25, 6, 20	7.15	46	8.20	3.96	3.02
Left	-24, 19, 8	5.97	14	5.87	4.22	2.52
Right	22, 8, 20	6.32	15	4.18	3.74	5.21
Right	20, 13, 13	5.39	5	4.62	2.74	3.24

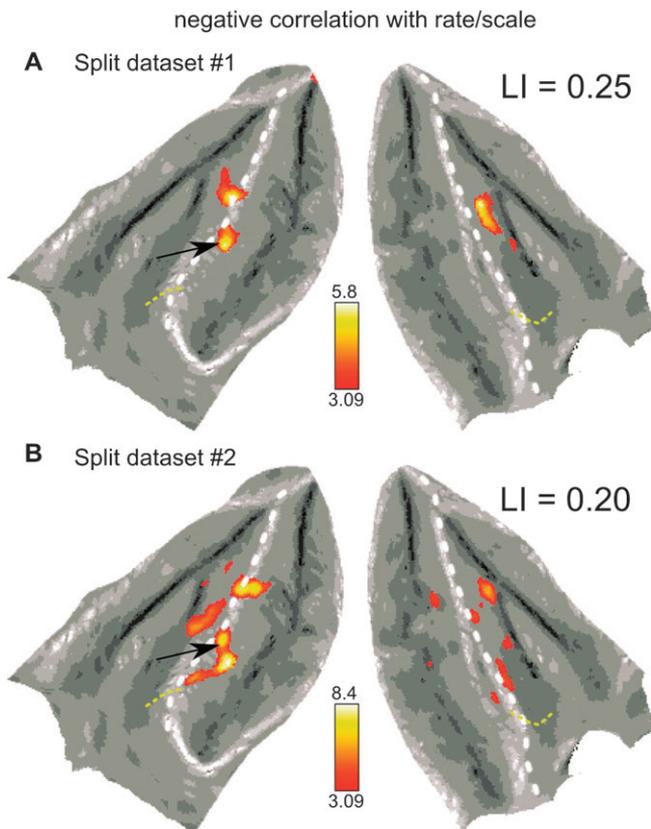


Figure 9. Within-experiment reliability for rate/scale analysis (group analysis, 3 monkeys). (A–B) SPM t -maps ($P < 0.001$ uncorrected, fixed effect) for rate/scale analysis (as shown Fig. 8A) for the 2 split data sets based on different scanning days.

For each subject, we also mapped (see Supplementary Fig. S3) the effect of scrambling within the sound class of human speech, which was the class having the most pronounced differences in rate/scale between intact and scrambled stimuli (Fig. 3). In all 3 monkeys, the scrambling effect reached significance, and the hemispheric difference of the scrambling effect (light green outlines in Supplementary Fig. S3) was significant in all 3 animals. Again, the pattern of these results was confirmed by the positive LI values (Table 2) in all 3 animals. Thus, the lack of consistency in the left lateralization of the general scrambling effect may, to a certain degree, reflect differences in the strength of the scrambling effect among the 3 classes of sound rather than a genuine subject difference. Indeed, the scrambling effect for human speech was both the most robust (Fig. 7) and the most consistently left lateralized (see Supplementary Fig. S3).

Supplementary Figure S4 shows the individual data for the regression with the rate/scale index. The activation pattern displayed a clearly left-sided dominance in all animals. In each

animal, an activation site in the STG, corresponding to the middle site of the group activation (Fig. 6A; Table 1), was present. The left lateralization of the regression was confirmed by the positive LI values (Table 2) in all 3 animals. The LI was positive even in subject M18, reaching 0.37. This confirms that the left asymmetry in higher-order auditory cortex is related to the processing of speech-like spectrotemporal patterns and is consistent across animals.

Discussion

Cortical and Subcortical Auditory System

Our fMRI experiment revealed activations in several subcortical regions, including the IC. The MGB activations were also observed in the deoxyglucose study of Poremba et al. (2003). This study relied on a unilateral removal of the IC and section of the commissures to create a deaf control hemisphere. Hence, cochlear and IC activations were reported neither in that study nor in any that we know of. At the cortical level, the general auditory activation was widespread and consistent with earlier full-brain mapping of the cortical auditory system (Poremba et al. 2003). These authors observed, as we did, activation of IPL, lateral bank of IPS, anterior insula, and frontal regions outside the auditory cortex. A conspicuous difference is the activation of sensorimotor cortex at the level of the head representation in our study, perhaps due to stimulus delivery via earphones in direct contact with the head. Single-cell studies have reported auditory responses in ventrolateral prefrontal cortex, extending into the lateral bank of the lateral orbitofrontal sulcus cortex (for review, see Romanski and Averbeck 2009), in ventral premotor cortex (Kohler et al. 2002), in posterior insula (Remedios et al. 2009), and in LIP (Mazzoni et al. 1996; Grunewald et al. 1999). These latter authors showed that LIP responses to noise bursts were induced by training. On the other hand, ventrolateral prefrontal and insular neurons were highly selective for monkey vocalizations (Romanski et al. 2005; Remedios et al. 2009). Finally, connections with auditory temporal regions have been reported for dorsal prefrontal cortex (Romanski et al. 1999) and primary visual cortex (Falchier et al. 2002).

Our main auditory activation reached its maximum in the right auditory core and was significantly greater in right area A1 than in the left. This rightward lateralization most likely corresponds to the posterior right hemispheric preference reported by Poremba et al. (2004) at the peak of activity along the STG (see their Fig. 2) and replicated in Gil-da-Costa et al. (2006) (their Supplementary Fig. S1). These 2 metabolic studies reported increases of 10% in the right hemisphere compared with the left, less than the 35% increase in MR signal we observed in the local maxima (Fig. 4B). In all 3 studies, this rightward asymmetry was observed for a broad range of stimuli suggesting that it reflects the processing of some low-level auditory feature common to all these stimuli. This is consistent with the observation that the asymmetry was present to some degree in subcortical structures, such as the MGB, and therefore this asymmetry may arise from subcortical regions.

Left Lateralization in Lateral Belt and Parabelt

Intact vocalizations evoked stronger responses than the scrambled control in several regions of the lateral belt and parabelt. The peak activations for scrambling were found lateral (~5–6

Table 2
Weighted LI scores of the contrasts indicated for the individual monkeys

Subtraction/regressor Individual	All He – silence	All Mv – silence	All Hs – silence	He – sHe	Mv – sMv	Hs – sHs	Rate/ scale
M14	–0.1	–0.1	–0.1	0.1	0.1	0.1	0.42
M18	–0.4	–0.4	–0.5	0	–0.1	0.1	0.37
M13	–0.1	–0.1	–0.2	0	0.1	0.3	0.17

Note: Positive LI values indicate left bias, negative values right bias.

mm) and slightly anterior (~1–2 mm) to the maxima for the main auditory response. Because of the location of its peak (Fig. 5B), we tentatively attribute the activation to region ML of the lateral belt even though the activation extended into neighboring regions, particularly into area AL. To a lesser degree, another cortical region located in the left anterior LS showed preferences for intact versus scrambled sounds. Since this site is located at the border between the medial belt and the auditory core, the attribution of this cluster to a precise cortical area remains difficult. Moreover, looking at individual sound classes, this anterior cluster was significant only for He (Fig. 7A).

Interestingly, the scrambling effect was left lateralized, as compared with the right lateralization of the main auditory activation. Hence, it is unlikely that the lateralization of the scrambling is due to asymmetric sound delivery because the lateralization for the main auditory activation was of opposite sign. The effect of scrambling was observed in the group and 2 of the 3 animals tested, meeting our minimum criterion (Nelissen et al. 2006) for a significant effect. Yet, unlike the rightward asymmetry, the scrambling effect and its lateralization depended on the type of sound. Lateralization was clearest using human speech, for which it was significant in all 3 monkeys tested. We suggest that the effectiveness of human speech in this regard is due to the combination of 2 factors: 1) Complex spectrotemporal processing is left biased in the rhesus monkey and 2) Among the stimuli used in this study, human speech shows the most complex spectrotemporal structure, thus the greatest scrambling effect.

A direct mapping of the effect of the rate/scale index revealed the involvement of several regions in the left belt and parabelt. The left asymmetry of this activation pattern was extremely robust, reaching significance in all 3 animals and, unlike the scrambling effect, did not depend on the degree of right lateralization of the main auditory activation. The most posterior region we observed is located near the AL–ML border, slightly anterior to the site of maximum scrambling effect. Neurons in these regions are selective for the slope and sign of frequency-modulated (FM) sweeps (Tian and Rauschecker 2004). Interestingly, an alternative interpretation of the rate/scale index is that of a generalized FM rate measure. Since the analysis of FM can be considered a first step in the processing of spectrotemporally complex sounds, it is not surprising that the FM-selective neurons could represent the first step in the analysis of complex spectrotemporal acoustic patterns (Rauschecker and Scott 2009). Indeed, Rauschecker et al. (1995) reported neurons in the lateral belt responsive to monkey vocalizations. Since this AL–ML region also is influenced by the scrambling, it is tempting to conclude that the scrambling affects the FM-selective neurons. Tian and Rauschecker (2004) have measured the optimal FM rate for ML and AL neurons with simple frequency sweeps. It is difficult, however, to predict from this physiological study the effect of scrambling on the complex stimuli used here because of

the different nature of the stimuli and the difference in response measures. It should nevertheless be noted that scrambling also disrupts pitch, so a scrambling effect on pitch-selective neurons (Bendor and Wang 2005) cannot be excluded.

In addition to the more posterior ML–AL region, 2 other regions in the middle STG also appeared in the rate/scale regression (Fig. 8; Table 1). The neuronal operation performed in these RPB regions, which belong to the ventral auditory pathway, is unknown, but neurons responsive to calls have been reported in these regions (Rauschecker et al. 1995; Tian et al. 2001; Russ et al. 2008), and it has been postulated that they receive their input from FM-selective neurons (Fig. 2 in Rauschecker and Scott 2009). The RPB is known (Kaas and Hackett 2000) to project to the upper bank of STS, and neurons integrating face and mouth movements with calls have been reported in this region by Ghazanfar et al. (2010). While a complete physiological identification of the middle STG site must await further investigation, it is worth stressing that its left asymmetric activation was robust insofar as it was observed with 2 different types of analysis, reliable since it was observed in the split data analysis, and consistent because it was present in all 3 subjects tested. It should be noted that an anatomical leftward interhemispheric asymmetry has been recently reported for monkey area Tpt (Gannon et al. 2008). This area is located at the level of the lower lip of the posterior third of the LS, too posterior to correspond to our left-lateralized activation sites (Fig. 5B).

Relationship with Earlier Imaging Studies

In comparing the results from PET and fMRI studies, one must take into account the major differences between both techniques. For instance, a lack of MR signal in the temporal regions and other regions near large cavities may explain the apparent discrepancy between fMRI and PET imaging. Notwithstanding these technical considerations, the relationship between our results and those of Poremba et al. (2004) is unclear. These authors showed that anterior temporal regions respond equally well to human speech and monkey vocalizations, as we observed in more posterior regions. Yet, they observed a left lateralization for monkey calls but not for human speech. Their finding was not replicated by Gil-da-Costa et al. (2006), who tested coos and screams separately, but it may be that using a single type of calls lacks the power of a variety of calls (Ghazanfar and Miller 2006). It should also be noted that our animals were required to maintain fixation in a window, while monkeys were free-viewing in the PET imaging studies. Furthermore, we used sequences of 2 different calls that are not spontaneously uttered by macaques (see Materials and Methods). Although, it is reasonable to assume that this has little effect on the early stage of the processing of vocalizations, further studies should investigate the different categories of calls to enable a straightforward comparison with the PET results of Gil-da-Costa et al. (2004) showing different responses elicited by screams and by coo calls. In our study, the left lateralization was observed for the scrambling effect and the regression with rate/scale but not for response level. The lack of a direct comparison with other conditions and the use of large regions of interest may have prevented Poremba et al. (2004) from observing the localized asymmetries we observed in the lateral belt and STG. The voice region described by Petkov et al. (2008) is located between the anterior pole region for which Poremba et al. (2004) described a left lateralization

and the most anterior STG region involved in the processing of rate/scale in the present study. While we cannot completely exclude the possibility that we may have missed this region because of lack of sensitivity of our measurements in particular in the temporal pole, an obvious question suggested by the functional profiles of our activation sites concerns how well this voice region and the temporal pole respond to human speech. This was not tested in the Petkov et al. (2008) study, nor in the other studies that reported brain areas involved in species-specific vocalizations (Gil-da-Costa et al. 2004, 2006). Our results show that many areas activated by conspecific calls can also show sensitivity to human (heterospecific) vocalizations. Indeed, it has been known for some time that even A1 neurons respond very well to human speech (Steinschneider et al. 1994). Hence, further work is required to investigate whether the brain areas reported in these earlier studies specifically process conspecific calls.

Relationship with Lesion Studies

The present results are in excellent agreement with a series of behavioral/lesion studies, despite the fact that these used Japanese macaques whereas we have used rhesus monkeys and concatenated 2 sounds. First, the coo calls used in the lesion study (Heffner HE and Heffner RS 1984) are spectrotemporally complex sounds, and the left asymmetry we observed depended on spectrotemporal complexity. This is also consistent with the left lateralization reported in lesion studies using human speech as stimuli (Dewson et al. 1969, 1970; Cowey and Dewson 1972). Second, the effects of unilateral left auditory cortex ablation were transient, while bilateral ablations were far more long-lasting (Heffner HE and Heffner RS 1986). This fits with our observation that the scrambling effect and the regression with rate/scale yield a bilateral activation with a leftward asymmetry. Indeed, this suggests that the right hemisphere can take over when the left side is damaged, exactly as shown by the lesion studies. Third, the unilateral left auditory cortex ablation yielded a consistent impairment in all 5 monkeys (Fig. 11 in Heffner HE and Heffner RS 1986): On average, monkeys needed 10 sessions to recover preoperative levels. However, individual differences were also observed: The time to recovery ranged from 5 to 14 sessions. This pattern is reminiscent of the variability in the left lateralization that we observed. For some measures, such as scrambling for human speech or regression with rate/scale, however, the lateralization effects were consistent across all 3 animals. Yet for other sounds, the effects were more variable (Table 2). Our study suggests 2 possible sources for this variability: differences in spectrotemporal complexity of the test stimuli and individual variations in the rightward asymmetry of the auditory core. Fourth, in a subsequent study using smaller lesions, Heffner HE and Heffner RS (1989) localized their effect to the posterior two-thirds of the STG, in good agreement with the regions showing a left asymmetry in the present study. Finally, the study of Harrington et al. (2001) linked the behavioral impairment in coo-call discrimination following bilateral ablations of auditory cortex to FM-selective mechanisms. Harrington et al. (2001) tested FM because of the role that the location of the inflection point between increasing and decreasing FM appeared to play in coo discrimination (May et al. 1988, 1989). As already mentioned, rate/scale tracks FM in complex sounds, and it yielded the largest and most robust

lateralization. From our results, we would predict that using other calls such as shrill barks in a similar lesion study would yield far less asymmetry. Indeed, our results, including those for human speech, clearly indicate that the left hemispheric dominance reflects the acoustic features of the vocalizations and not their general nature.

Evolution of Vocal Communication in Primates

While lateralization of the processing of conspecific vocalizations has been reported for other species (e.g., Ehret 1987; Wetzel et al. 1998), such lateralization in nonhuman primates is directly relevant to understanding the human processing of speech because of the evolutionary proximity of monkeys to humans. In the present study, we did find a consistent left hemispheric bias for speech-like spectrotemporal complex patterns in the rhesus monkey. Hence, we can minimally speculate that, as speech evolved in hominids, brain areas that were already suited for complex spectrotemporal processing were naturally recruited for speech processing, yielding a left hemispheric dominance. The left dominance in the monkey extends into the RPB, a high-level auditory cortical area in the anterior what stream (Hackett et al. 1999), consistent with the high-level localization of left dominance in human speech processing (see Introduction). This adds to the growing list of monkey cortical regions that were prepared during the course of evolution for the advent of speech communication: the STS by virtue of its sensitivity to slow visual modulations (Ghazanfar et al. 2010), the ventral premotor cortex by its association of visual and auditory signals with motor signals related to mouth and hand actions (Kohler et al. 2002; Ferrari et al. 2003), and the ventral prefrontal cortex by its integration of faces and corresponding vocalizations (Sugihara et al. 2006). Thus, owing to their capacity for processing spectrotemporally complex sounds, the belt and RPB of monkey, especially in the left hemisphere, were ready for the advent of speech communication.

Supplementary Material

Supplementary material, Table S1, and Figures S1–S4 can be found at: <http://www.cercor.oxfordjournals.org/>

Funding

European Union grants—Sensoprim (MEST-CT-2004-007825); Neurocom (NEST 012738); EF 05/14 from the KU Leuven Research council; G 151.04 from the Fonds voor Wetenschappelijk onderzoek (FWO) to G.A.O. O.J. was a doctoral fellow supported by Sensoprim.

Notes

The help of P. Kayenbergh, G. Meulemans, M. Depaep, C. Fransen, A. Coeman, C. Giffard, and S. Kovacs is kindly acknowledged. The authors are indebted to M. Hauser for supplying the macaque vocalizations and for comments on earlier versions of the manuscript as well as to C. Pallier, S. Raiguel, and S. Shamma for helpful comments. Sinerem was kindly provided by Guerbet (Roissy, France). *Conflict of Interest:* None declared.

References

Avants B, Epstein C, Grossman M, Gee J. 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated

- labeling of elderly and neurodegenerative brain. *Med Image Anal.* 12:26–41.
- Averbeck BB, Romanski LM. 2006. Probabilistic encoding of vocalizations in macaque ventral lateral prefrontal cortex. *J Neurosci.* 26:11023–11033.
- Baumgart F, Kaulisch T, Tempelmann C, Gaschler-Markefski B, Tegeler C, Schindler F, Stiller D, Scheich H. 1998. Electrodynamic headphones and woofers for application in magnetic resonance imaging scanners. *Med Phys.* 25:2068–2070.
- Belin P, Fecteau S, Charest I, Nicastrò N, Hauser MD, Armony JL. 2008. Human cerebral response to animal affective vocalizations. *Proc Biol Sci.* 275:473–481.
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. 2000. Voice-selective areas in human auditory cortex. *Nature.* 403:309–312.
- Bendor D, Wang X. 2005. The neuronal representation of pitch in primate auditory cortex. *Nature.* 436:1161–1165.
- Bradbury JW, Vehrencamp SL. 1998. *Principles of animal communication.* Oxford: Blackwell.
- Chi T, Ru P, Shamma SA. 2005. Multiresolution spectrotemporal analysis of complex sounds. *J Acoust Soc Am.* 118:887–906.
- Cowey A, Dewson JH. 1972. Effects of unilateral ablation of superior temporal cortex on auditory sequence discrimination in *Macaca mulatta*. *Neuropsychologia.* 10:279–289.
- Dehaene-Lambertz G, Pallier C, Serniclaes W, Sprenger-Charolles L, Jobert A, Dehaene S. 2005. Neural correlates of switching from auditory to speech perception. *Neuroimage.* 24:21–33.
- Depireux DA, Simon JZ, Klein DJ, Shamma SA. 2001. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J Neurophysiol.* 85:1220–1234.
- Dewson JH, Cowey A, Weiskrantz L. 1970. Disruptions of auditory sequence discrimination by unilateral and bilateral cortical ablations of superior temporal gyrus in the monkey. *Exp Neurol.* 28:529–548.
- Dewson JH, Pribram KH, Lynch JC. 1969. Effects of ablations of temporal cortex upon speech sound discrimination in the monkey. *Exp Neurol.* 24:579–591.
- Dhanjal NS, Handunnetthi L, Patel MC, Wise RJS. 2008. Perceptual systems controlling speech production. *J Neurosci.* 28:9969–9975.
- Ehret G. 1987. Left hemisphere advantage in the mouse brain for recognizing ultrasonic communication calls. *Nature.* 325:249–251.
- Ekstrom LB, Roelfsema PR, Arsenault JT, Bonmassar G, Vanduffel W. 2008. Bottom-up dependent gating of frontal signals in early visual cortex. *Science.* 321:414–417.
- Falchier A, Clavagnier S, Barone P, Kennedy H. 2002. Anatomical evidence of multimodal integration in primate striate cortex. *J Neurosci.* 22:5749–5759.
- Ferrari PF, Gallese V, Rizzolatti G, Fogassi L. 2003. Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur J Neurosci.* 17:1703–1714.
- Fischer J, Teufel C, Drolet M, Patzelt A, Rübsemann R, von Cramon DY, Schubotz RI. 2009. Orienting asymmetries and lateralized processing of sounds in humans. *BMC Neurosci.* 10:14.
- Gannon PJ, Kheck N, Hof PR. 2008. Leftward interhemispheric asymmetry of macaque monkey temporal lobe language area homolog is evident at the cytoarchitectural, but not gross anatomic level. *Brain Res.* 1199:62–73.
- Georgieva S, Peeters R, Kolster H, Todd JT, Orban GA. 2009. The processing of three-dimensional shape from disparity in the human brain. *J Neurosci.* 29:727–742.
- Ghazanfar AA, Chandrasekaran C, Logothetis NK. 2008. Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *J Neurosci.* 28:4457–4469.
- Ghazanfar AA, Chandrasekaran C, Morrill RJ. 2010. Dynamic, rhythmic facial expressions and the superior temporal sulcus of macaque monkeys: implications for the evolution of audiovisual speech. *Eur J Neurosci.* 31:1807–1817.
- Ghazanfar AA, Miller CT. 2006. Language evolution: loquacious monkey brains? *Curr Biol.* 16:R879–R881.
- Ghazanfar AA, Smith-Rohrberg D, Hauser MD. 2001. The role of temporal cues in rhesus monkey vocal recognition: orienting asymmetries to reversed calls. *Brain Behav Evol.* 58:163–172.
- Gifford GW, Hauser MD, Cohen YE. 2003. Discrimination of functionally referential calls by laboratory-housed rhesus macaques: implications for neuroethological studies. *Brain Behav Evol.* 61:213–224.
- Gil-da-Costa R, Braun A, Lopes M, Hauser MD, Carson RE, Herscovitch P, Martin A. 2004. Toward an evolutionary perspective on conceptual representation: species-specific calls activate visual and affective processing systems in the macaque. *Proc Natl Acad Sci U S A.* 101:17516–17521.
- Gil-da-Costa R, Hauser MD. 2006. Vervet monkeys and humans show brain asymmetries for processing conspecific vocalizations, but with opposite patterns of laterality. *Proc Biol Sci.* 273:2313–2318.
- Gil-da-Costa R, Martin A, Lopes MA, Muñoz M, Fritz JB, Braun AR. 2006. Species-specific calls activate homologs of broca's and wernicke's areas in the macaque. *Nat Neurosci.* 9:1064–1070.
- Grunewald A, Linden JF, Andersen RA. 1999. Responses to auditory stimuli in macaque lateral intraparietal area. I. Effects of training. *J Neurophysiol.* 82:330–342.
- 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.
- Hackett TA, Stepniewska I, Kaas JH. 1998. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *J Comp Neurol.* 394:475–495.
- Hackett TA, Stepniewska I, Kaas JH. 1999. Prefrontal connections of the parabelt auditory cortex in macaque monkeys. *Brain Res.* 817:45–58.
- Harrington IA, Heffner RS, Heffner HE. 2001. An investigation of sensory deficits underlying the aphasia-like behavior of macaques with auditory cortex lesions. *Neuroreport.* 12:1217–1221.
- Hauser MD. 1996. *The evolution of communication.* Cambridge (MA): The MIT Press.
- Hauser MD, Andersson K. 1994. Left hemisphere dominance for processing vocalizations in adult, but not infant, rhesus monkeys: field experiments. *Proc Natl Acad Sci U S A.* 91:3946–3948.
- Heffner HE, Heffner RS. 1984. Temporal lobe lesions and perception of species-specific vocalizations by macaques. *Science.* 226:75–76.
- Heffner HE, Heffner RS. 1986. Hearing loss in Japanese macaques following bilateral auditory cortex lesions. *J Neurophysiol.* 55:256–271.
- Heffner HE, Heffner RS. 1989. Effect of restricted cortical lesions on absolute thresholds and aphasia-like deficits in Japanese macaques. *Behav Neurosci.* 103:158–169.
- Jesteadt W, Wier CC, Green DM. 1977. Intensity discrimination as a function of frequency and sensation level. *J Acoust Soc Am.* 61:169–177.
- Joly O, Vanduffel W, Orban GA. 2009. The monkey ventral premotor cortex processes 3D shape from disparity. *Neuroimage.* 47:262–272.
- Kaas JH, Hackett TA. 2000. Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci U S A.* 97:11793–11799.
- Kayser C, Logothetis NK. 2009. Directed interactions between auditory and superior temporal cortices and their role in sensory integration. *Front Integr Neurosci.* 3:7.
- Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, Christensen GE, Collins DL, Gee J, Hellier P, et al. 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage.* 46:786–802.
- Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G. 2002. Hearing sounds, understanding actions: action representation in mirror neurons. *Science.* 297:846–848.
- Kovacs S, Peeters R, Smits M, De Ridder D, Van Hecke P, Sunaert S. 2006. Activation of cortical and subcortical auditory structures at 3 T by means of a functional magnetic resonance imaging paradigm suitable for clinical use. *Invest Radiol.* 41:87–96.
- Leite FP, Tsao D, Vanduffel W, Fize D, Sasaki Y, Wald LL, Dale AM, Kwong KK, Orban GA, Rosen BR, et al. 2002. Repeated fMRI using iron oxide contrast agent in awake, behaving macaques at 3 Tesla. *Neuroimage.* 16:283–294.

- May B, Moody D, Stebbins W. 1988. The significant features of Japanese macaque coo sounds: a psychophysical study. *Anim Behav.* 36:1432-1444.
- May B, Moody D, Stebbins W. 1989. Categorical perception of conspecific communication sounds by Japanese macaques, *Macaca fuscata*. *J Acoust Soc Am.* 85:837.
- Mazzoni P, Bracewell RM, Barash S, Andersen RA. 1996. Spatially tuned auditory responses in area lip of macaques performing delayed memory saccades to acoustic targets. *J Neurophysiol.* 75:1233-1241.
- McLaren DG, Kosmatka KJ, Oakes TR, Kroenke CD, Kohama SG, Matochik JA, Ingram DK, Johnson SC. 2009. A population-average MRI-based atlas collection of the rhesus macaque. *Neuroimage.* 45:52-59.
- Narain C, Scott SK, Wise RJS, Rosen S, Leff A, Iversen SD, Matthews PM. 2003. Defining a left-lateralized response specific to intelligible speech using fMRI. *Cereb Cortex.* 13:1362-1368.
- Nelissen K, Vanduffel W, Orban GA. 2006. Charting the lower superior temporal region, a new motion-sensitive region in monkey superior temporal sulcus. *J Neurosci.* 26:5929-5947.
- Patterson RD, Allerhand MH, Giguère C. 1995. Time-domain modeling of peripheral auditory processing: a modular architecture and a software platform. *J Acoust Soc Am.* 98:1890-1894.
- Petersen MR, Beecher MD, Zoloth SR, Moody DB, Stebbins WC. 1978. Neural lateralization of species-specific vocalizations by Japanese macaques (*Macaca fuscata*). *Science.* 202:324-327.
- Petkov CI, Kayser C, Augath M, Logothetis NK. 2006. Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biol.* 4:e215.
- Petkov CI, Kayser C, Steudel T, Whittingstall K, Augath M, Logothetis NK. 2008. A voice region in the monkey brain. *Nat Neurosci.* 11:367-374.
- Poeppl D, Guillemin A, Thompson J, Fritz J, Bavelier D, Braun AR. 2004. Auditory lexical decision, categorical perception, and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia.* 42:183-200.
- Poremba A, Malloy M, Saunders RC, Carson RE, Herscovitch P, Mishkin M. 2004. Species-specific calls evoke asymmetric activity in the monkey's temporal poles. *Nature.* 427:448-451.
- Poremba A, Saunders RC, Crane AM, Cook M, Sokoloff L, Mishkin M. 2003. Functional mapping of the primate auditory system. *Science.* 299:568-572.
- Rauschecker JP. 1998. Parallel processing in the auditory cortex of primates. *Audiol Neurootol.* 3(2-3):86-103.
- Rauschecker JP, Scott SK. 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci.* 12:718-724.
- Rauschecker JP, Tian B. 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc Natl Acad Sci U S A.* 97:11800-11806.
- Rauschecker JP, Tian B, Hauser M. 1995. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science.* 268:1111-1114.
- Remedios R, Logothetis NK, Kayser C. 2009. An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. *J Neurosci.* 29:1034-1045.
- Romanski LM, Averbeck BB. 2009. The primate cortical auditory system and neural representation of conspecific vocalizations. *Annu Rev Neurosci.* 32:315-346.
- Romanski LM, Averbeck BB, Diltz M. 2005. Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. *J Neurophysiol.* 93:734-747.
- Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat Neurosci.* 2:1131-1136.
- Ruggiero MA, Temchin AN. 2005. Unexceptional sharpness of frequency tuning in the human cochlea. *Proc Natl Acad Sci U S A.* 102:18614-18619.
- Russ BE, Ackelson AL, Baker AE, Cohen YE. 2008. Coding of auditory-stimulus identity in the auditory non-spatial processing stream. *J Neurophysiol.* 99:87-95.
- Saleem K, Logothetis N. 2006. A combined MRI and histology Atlas of the Rhesus monkey brain. San Diego (CA): Academic Press.
- Samson F, Zeffiro TA, Toussaint A, Belin P. 2011. Stimulus complexity and categorical effects in human auditory cortex: an activation likelihood estimation meta-analysis. *Front. Psychology.* 1:241.
- Schroeder CE, Lakatos P, Kajikawa Y, Partan S, Puce A. 2008. Neuronal oscillations and visual amplification of speech. *Trends Cogn Sci.* 12:106-113.
- Scott SK, Blank CC, Rosen S, Wise RJ. 2000. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain.* 123(Pt 12):2400-2406.
- Serafini JV, Moody DB, Stebbins WC. 1982. Frequency selectivity of the monkey's auditory system: psychophysical tuning curves. *J Acoust Soc Am.* 71:1513-1518.
- Spitsyna G, Warren JE, Scott SK, Turkheimer FE, Wise RJS. 2006. Converging language streams in the human temporal lobe. *J Neurosci.* 26:7328-7336.
- Steinschneider M, Schroeder CE, Arezzo JC, Vaughan HG. 1994. Speech-evoked activity in primary auditory cortex: effects of voice onset time. *Electroencephalogr Clin Neurophysiol.* 92:30-43.
- Sugihara T, Diltz MD, Averbeck BB, Romanski LM. 2006. Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. *J Neurosci.* 26:11138-11147.
- Teufel C, Ghazanfar AA, Fischer J. 2010. On the relationship between lateralized brain function and orienting asymmetries. *Behav Neurosci.* 124:437-445.
- Tian B, Rauschecker JP. 2004. Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol.* 92:2993-3013.
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. 2001. Functional specialization in rhesus monkey auditory cortex. *Science.* 292:290-293.
- Vanduffel W, Fize D, Mandeville JB, Nelissen K, Hecke PV, Rosen BR, Tootell RB, Orban GA. 2001. Visual motion processing investigated using contrast agent-enhanced fMRI in awake behaving monkeys. *Neuron.* 32:565-577.
- Wetzel W, Wagner T, Ohl FW, Scheich H. 1998. Categorical discrimination of direction in frequency-modulated tones by mongolian gerbils. *Behav Brain Res.* 91:29-39.
- Wilke M, Lidzba K. 2007. LI-tool: a new toolbox to assess lateralization in functional MR-data. *J Neurosci Methods.* 163:128-136.