

Functional MRI of the vocalization-processing network in the macaque brain

Michael Ortiz-Rios^{1,2,3*}, Paweł Kuśmierk¹, Iain DeWitt¹, Denis Archakov^{1,4}, Frederico A. C. Azevedo^{2,3}, Mikko Sams⁴, Iiro P. Jääskeläinen⁴, Georgios A. Keliris^{2,5,6} and Josef P. Rauschecker^{1,4,7*}

¹ Department of Neuroscience, Georgetown University Medical Center, Washington, DC, USA, ² Department of Physiology of Cognitive Processes, Max Planck Institute for Biological Cybernetics, Tübingen, Germany, ³ IMPRS for Cognitive and Systems Neuroscience, Tübingen, Germany, ⁴ Brain and Mind Laboratory, Department of Neuroscience and Biomedical Engineering, Aalto University School of Science, Aalto, Finland, ⁵ Bernstein Centre for Computational Neuroscience, Tübingen, Germany, ⁶ Department of Biomedical Sciences, University of Antwerp, Wilrijk, Belgium, ⁷ Institute for Advanced Study and Department of Neurology, Klinikum Rechts der Isar, Technische Universität München, München, Germany

OPEN ACCESS

Edited by:

Monica Munoz-Lopez,
University of Castilla-La Mancha,
Spain

Reviewed by:

Simon Baumann,
Newcastle University, UK
Huan Luo,
Chinese Academy of Sciences, China
Olivier Joly,
MRC Cognition and Brain Sciences
Unit, UK

*Correspondence:

Michael Ortiz-Rios and
Josef P. Rauschecker,
Department of Neuroscience,
Georgetown University Medical
Center, NRB WP19, 3970 Reservoir
Rd. NW, Washington, DC 20057, USA
michael.ortiz@tuebingen.mpg.de;
rauschej@georgetown.edu

Specialty section:

This article was submitted to Auditory Cognitive Neuroscience, a section of the journal *Frontiers in Neuroscience*

Received: 30 December 2014

Accepted: 17 March 2015

Published: 01 April 2015

Citation:

Ortiz-Rios M, Kuśmierk P, DeWitt I, Archakov D, Azevedo FAC, Sams M, Jääskeläinen IP, Keliris GA and Rauschecker JP (2015) Functional MRI of the vocalization-processing network in the macaque brain. *Front. Neurosci.* 9:113. doi: 10.3389/fnins.2015.00113

Using functional magnetic resonance imaging in awake behaving monkeys we investigated how species-specific vocalizations are represented in auditory and auditory-related regions of the macaque brain. We found clusters of active voxels along the ascending auditory pathway that responded to various types of complex sounds: inferior colliculus (IC), medial geniculate nucleus (MGN), auditory core, belt, and parabelt cortex, and other parts of the superior temporal gyrus (STG) and sulcus (STS). Regions sensitive to monkey calls were most prevalent in the anterior STG, but some clusters were also found in frontal and parietal cortex on the basis of comparisons between responses to calls and environmental sounds. Surprisingly, we found that spectrotemporal control sounds derived from the monkey calls (“scrambled calls”) also activated the parietal and frontal regions. Taken together, our results demonstrate that species-specific vocalizations in rhesus monkeys activate preferentially the auditory ventral stream, and in particular areas of the antero-lateral belt and parabelt.

Keywords: auditory cortex, monkey, species-specific calls, spectrotemporal features, higher-level representations

Introduction

The concept of two streams in auditory cortical processing, analogous to that in visual cortex (Mishkin et al., 1983), was proposed more than a decade ago (Rauschecker, 1998a; Rauschecker and Tian, 2000). The concept was supported by contrasting patterns of anatomical connections in the macaque from anterior/ventral and posterior/dorsal belt regions of auditory cortex to segregated domains of lateral prefrontal cortex (Romanski et al., 1999) and by different physiological properties of these belt regions. In particular, the anterior lateral belt (area AL) in the macaque exhibited enhanced selectivity for the identity of sounds (monkey vocalizations), whereas the caudal lateral belt (area CL) was particularly selective to sound location (Tian et al., 2001; see also Kuśmierk and Rauschecker, 2014). Evidence for segregated streams of auditory cortical processing has also been provided in human studies (Maeder et al., 2001; Arnott et al., 2004; Ahveninen et al., 2006).

Use of species-specific vocalizations for auditory stimulation in the macaque is of particular interest in the context of the ongoing debate about the evolution of speech and language (Rauschecker, 2012; Bornkessel-Schlesewsky et al., 2015). Comparative approaches have focused

on identifying the common neural networks involved in the processing of speech in humans and of vocalizations in non-human primates (Gil-da-Costa et al., 2004; Frey et al., 2008, 2014; Petrides and Pandya, 2009; Joly et al., 2012b). Monkey calls convey semantic information about objects and events in the environment as well as about affective states of individuals, similar to information contained in human communication sounds and speech (Cheney and Seyfarth, 1990; Ghazanfar and Hauser, 1999; Yovel and Belin, 2013). An open question regarding the vocalization-processing network in the macaque brain is whether it also carries information about the motor actions necessary to produce the vocalizations, as has been shown in humans listening to speech and music (Wilson et al., 2004; Leaver et al., 2009).

Several studies have examined the representation of complex sounds, including vocalizations, in the macaque brain using neuroimaging techniques (Poremba et al., 2003; Petkov et al., 2008; Joly et al., 2012b). In particular, the first fMRI study by Petkov et al. (2008) found activation specific to monkey vocalizations in the anterior STG region. One of the aims in later studies has been to characterize the physiological properties of the anterior superior temporal (aSTG) region that shows sensitivity to higher-level spectrotemporal features in vocalizations (Russ et al., 2008; Kikuchi et al., 2010, 2014; Perrodin et al., 2011; Fukushima et al., 2014). A recent comparative study by Joly et al. (2012b) replicated and extended these results by analyzing fMRI images of the entire brain and found an involvement of orbitofrontal cortex in the processing of monkey vocalizations. Given that the ventral pathway continues into orbitofrontal and ventrolateral prefrontal cortex (vlPFC) (Barbas, 1993; Romanski et al., 1999; Cohen et al., 2007; Petkov et al., 2015), this finding is of particular interest.

In humans, the ventral auditory pathway is thought to be particularly involved in the recognition and identification of vocalizations as well as speech (Binder et al., 2000; DeWitt and Rauschecker, 2012). By contrast, the dorsal pathway is involved primarily in processing sound source location and motion in both humans and animals (Maeder et al., 2001; Tian et al., 2001; Arnott et al., 2004). However, a recent proposal, derived from both human and non-human primate studies, suggests that the dorsal stream may also play a role in sensorimotor integration and control of complex sounds, including speech (Rauschecker and Scott, 2009; Rauschecker, 2011). Thus, activation of frontal and parietal regions might also be expected when monkeys are presented with conspecific vocalization sounds.

Here we identified which brain regions of the macaque monkey are sensitive to conspecific vocalizations using whole-brain functional magnetic resonance imaging (fMRI). We found the most distinct activation in the anterior STG and along the auditory ventral stream, but some clusters of activation were also found in prefrontal, premotor, and parietal cortex when comparing monkey vocalizations to environmental sounds. These findings are discussed in terms of their functional significance.

Materials and Methods

Subjects

Two male rhesus monkeys (*Macaca mulatta*) weighing 10–12 kg participated in our awake-fMRI experiments. Each animal was

implanted with an MRI-compatible headpost (Applied Prototype) secured to the skull with ceramic screws (Thomas Recording), plastic strips, and bone cement (Osteobond, Zimmer). All surgical procedures were performed under general anesthesia with isoflurane (1–2%) following pre-anesthetic medication with ketamine (13 mg/kg) and midazolam (0.12 mg/kg). The experiments were approved by the Georgetown University Animal Care and Use Committee and conducted in accordance with standard NIH guidelines.

Behavioral Training

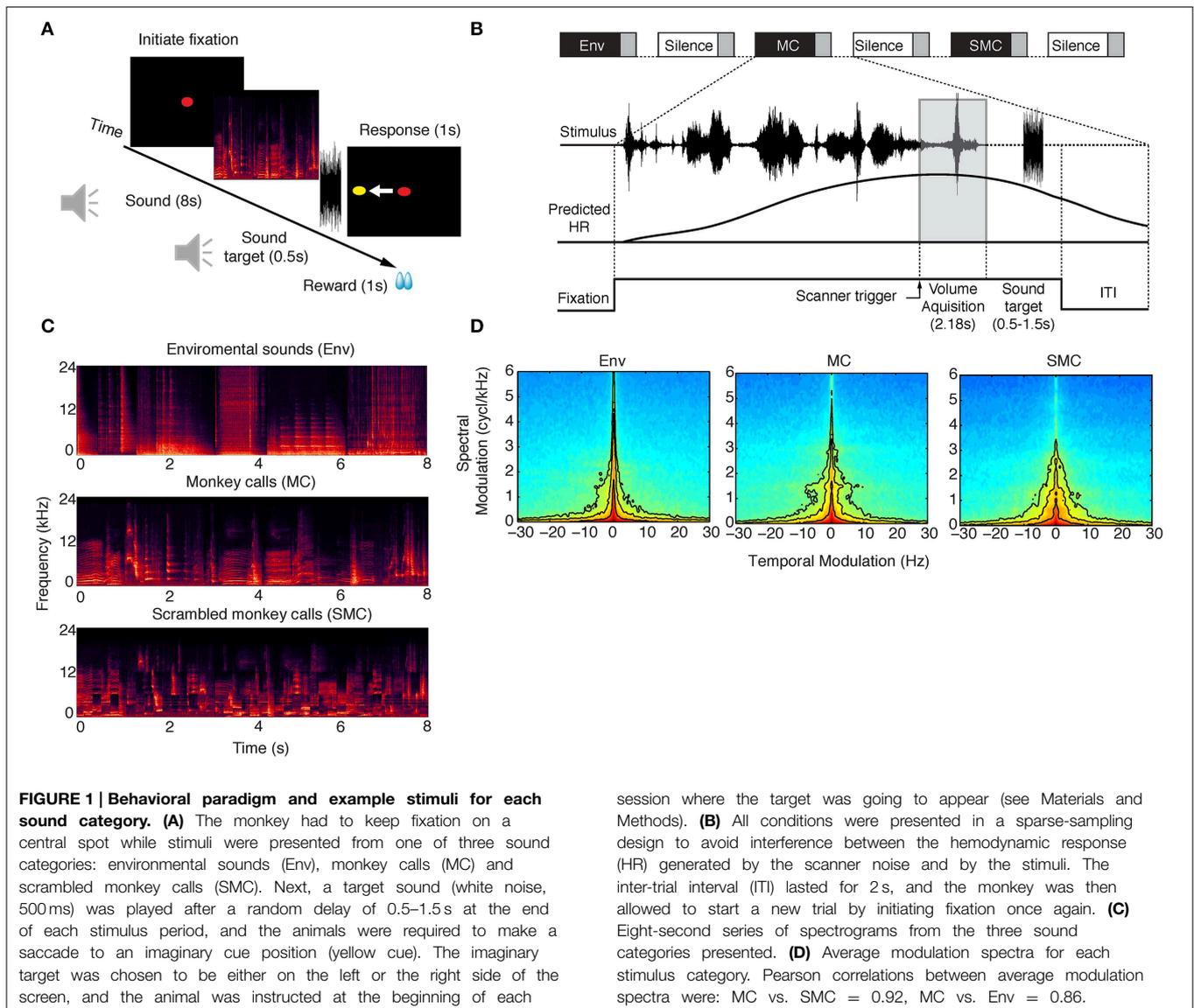
To ensure the monkeys attended to each stimulus for which a brain volume was acquired, we adapted a go/no-go auditory discrimination task (Kuśmierk and Rauschecker, 2009; Kikuchi et al., 2010) for sparse-sampling functional MRI.

First, each monkey was trained to lie in sphinx position in an MRI-compatible primate chair (Applied Prototype) placed inside a double-walled acoustic chamber simulating the scanner environment. Inside the chamber, the animals were trained to be accustomed to wearing headphone equipment and hearing (simulated) scanner noise, presented by a loudspeaker. Eye movements were monitored using an infrared eye-tracking system (ISCAN). Analog output of the tracker was sampled with an analog-to-digital conversion device (National Instruments). A PC running Presentation software (Neurobehavioral Systems) was used to present visual and auditory stimuli, control the reward system, and trigger imaging data acquisition (see below).

After the animal completed the fixation training, a go/no-go auditory discrimination task was introduced, in which the monkeys could initiate a trial by holding fixation on a central red spot while a block of auditory stimuli would be simultaneously presented. After the first 6 s of auditory stimulation, a trigger was sent to the scanner, starting the acquisition of an image volume (**Figure 1B**). Following acquisition and a random delay, the target sound (white noise) was presented, cueing a saccade to the left or to the right side as signaled at the beginning of each experimental session (**Figure 1A**). To provide feedback, after the response window, a yellow spot was shown indicating the correct target location. Finally, contingent on performance, the animal received a juice reward. An inter-trial interval of at least 2 s was enforced before the next trial could be initiated by fixation. Every sound presentation trial was followed by a “silence” trial, allowing for measurement of baseline blood oxygen level dependent (BOLD) signal. Monkey 1 (M1) performed the task correctly for over 90% of the trials. Monkey 2 (M2) was not able to perform the saccadic go/no-go discrimination task with high accuracy and was therefore scanned while passively listening to the acoustic stimuli. To ensure stable attention, M2 was rewarded for successfully holding fixation throughout the trial.

Auditory Stimuli

Three sound categories were used in the experiments: environmental sounds (Env), monkey vocalizations or calls (MC), and scrambled monkey calls (SMC). Spectrograms of example clips from each of these three categories are illustrated in **Figure 1C**. Environmental sounds were obtained from multiple



online sources and from recordings made in our laboratory facilities (Kuśmierk and Rauschecker, 2009). They included the sounds of vehicles, cages, water, food containers, clocks, cameras, applause, coins, footsteps, chewing, heartbeats, horns, and telephones ($n = 56$). The mean duration of the Env stimuli was 1.14 s (range: 0.96–2.6 s). Monkey calls were obtained from recordings made outside our colony [M. Hauser and/or Laboratory of Neuropsychology (LN) library]. Monkey vocalizations ($n = 63$) consisted of grunts, barks, warbles, coos, and screams, as used in prior studies (Rauschecker et al., 1995; Tian et al., 2001; Kuśmierk et al., 2012). The mean duration of the vocalization stimuli was 0.67 s (range: 0.13–2.34 s). SMC were generated by randomly rearranging 200 ms by 1-octave tiles of the constant-Q spectrogram (Brown, 1991) for each monkey call and reconstructing a time-domain waveform with an inverse transform (Schörkhuber and Klapuri, 2010). Transposition along the time axis was not constrained while transposition along the frequency

axis was restricted to displacement by a single octave. For each trial, a random selection of stimuli from one class (MC, Env, or SMC) was arranged sequentially into a smooth auditory clip that lasted for the duration of the trial (8 s).

Sounds were presented through modified electrostatic in-ear headphones (SRS-005S + SRM-252S, STAX), mounted on ear-mold impressions of each animal's pinna (Sarkey Eden Prairie) and covered with a custom-made earmuff system for sound attenuation. To match loudness, the stimuli were played through the sound presentation system and re-recorded with a probe microphone (Brüel and Kjær, type 4182 SPL meter) inserted in the ear-mold of an anesthetized monkey. The recordings were then filtered with an inverted macaque audiogram (Jackson et al., 1999) to simulate the effect of different ear sensitivity at different frequencies, analogous to the dB(A) scale for humans. The stimuli were finally equalized so that they produced equal maximum root mean square (RMS) amplitude (using

a 200-ms sliding window) in filtered recordings (Kuśmierk and Rauschecker, 2009). During experiments, all stimuli were amplified (Yamaha AX-496) and delivered at a calibrated RMS amplitude of ~ 80 dB SPL.

Analyses of Sound Categories

A modulation spectrum analysis (Singh and Theunissen, 2003) was performed for each sound with the STRFpak Matlab toolbox (<http://strfpak.berkeley.edu>). We obtained a spectrogram of each sound by decomposing it into frequency bands using a bank of Gaussian filters (244 bands, filter width = 125 Hz). The filters were evenly spaced on the frequency axis (64–48,000 Hz) and separated from each other by one standard deviation. The decomposition resulted in a set of narrow-band signals, which were then cross-correlated with each other and themselves to yield a cross-correlation matrix. This matrix was calculated for time delays of ± 150 ms, and the two-dimensional Fourier transform of this matrix was calculated to obtain the modulation spectrum of each sound (Figure 1D).

Data Acquisition

Images were acquired with a horizontal MAGNETOM Trio 3-T scanner (Siemens) with a 60-cm bore diameter. A 12-cm custom-made saddle shape radiofrequency coil (Windmiller Kolster Scientific) covered the entire brain and was optimized for imaging the temporal lobe. The time series consisted of gradient-echo echo-planar (GE-EPI) whole-brain images obtained in a sparse acquisition design. Sparse sampling allows single volumes to be recorded coincidentally with the predicted peak of the evoked hemodynamic response (Hall et al., 1999). This helps to avoid contamination of the measured stimulus-specific BOLD response by the scanner-noise-evoked BOLD response. Further, by triggering acquisition 6 s after stimulus onset, the auditory stimulus was presented without acoustic interference from gradient-switching noise, typical of a continuous fMRI design. For the functional data, individual volumes with 25 ordinal slices were acquired with an interleaved single-shot GE-EPI sequence (TE = 34 ms, TA = 2.18 s, flip angle = 90° , field of view (FOV) = 100×100 mm², matrix size = 66×66 voxels, slice thickness = 1.9 mm, voxel size = $1.5 \times 1.5 \times 1.9$ mm³). On each experiment day, a low-resolution FLASH anatomical scan was acquired with the same geometry as the functional images (TE = 14 ms, TR = 3 s, TA = 2.18 s, FOV = 100×100 mm², matrix = 512×512 voxels, slice thickness = 1.9 mm, number of averages = 2, flip angle = 150°). For overlaying our functional images, we created a high-resolution anatomical template ($0.5 \times 0.5 \times 0.5$ mm³ isotropic voxels) by averaging five high-resolution anatomical scans acquired under general anesthesia with an MP-RAGE sequence (TE = 3.0 ms, TR = 2.5 s, flip angle = 8° , FOV = $116 \times 96 \times 128$ mm³; matrix = $232 \times 192 \times 256$ voxels).

Data Analysis

For M1, nine EPI runs (180 time points each) were acquired over six sessions. For M2, seven runs were acquired over four sessions. All data analyses were performed using AFNI (Cox, 1996) (<http://afni.nimh.nih.gov/afni>), FreeSurfer (Dale et al., 1999; Fischl et al., 1999) (<http://surfer.nmr.mgh.harvard.edu/>),

SUMA (<http://afni.nimh.nih.gov/>) and custom code written in Matlab (MathWorks). Preprocessing involved slice timing correction, motion correction (relative to the run-specific mean GE-EPI), spatial smoothing with a 3.0 mm full width at half-maximum Gaussian kernel, and normalization of the time series at each voxel by its mean. All volumes that had motion values with shifts >0.5 mm and/or rotations $>0.5^\circ$ were excluded from further analyses. Lastly, we performed linear least-squares detrending to remove non-specific variations (i.e., scanner drift). Following preprocessing, data were submitted to generalized linear model analyses. The model included three stimulus-specific regressors and six estimated motion regressors of no interest. For each stimulus category (Env, MC, SMC) we estimated a regressor by convolving a one-parameter gamma distribution estimate of the hemodynamic response function with the square-wave stimulus function. We performed *t*-tests contrasting all sounds vs. baseline (“silence” trials), MC vs. Env, and MC vs. SMC. Finally we co-registered and normalized our functional data to the population-average MRI-based template for rhesus monkeys 112RM-SL (McLaren et al., 2009) and then displayed the results on a semi-inflated cortical surface of the template extracted with Freesurfer and displayed with SUMA to facilitate visualization and identification of cortical activations. The anatomical boundaries described here are based on the macaque brain atlas of Saleem and Logothetis (2012).

To quantify the lateralization of the BOLD response across hemispheres we measured a lateralization index [LI = $(R_h - L_h)/(R_h + L_h)$], where R_h and L_h are the mean responses in the right and left hemisphere, respectively. The LI curve analyses ensure that the lateralization effect is not caused by small numbers of highly activated voxels across hemispheres. The LI curves were based on the *t*-values obtained from each contrast condition and were calculated using the LI-toolbox (Wilke and Lidzba, 2007) with the following options: ± 5 mm mid-sagittal exclusive mask, clustering with a minimum of 5 voxels and default bootstrapping parameters (min/max sample size 5/10,000 and bootstrapping set to 25% of data). The bootstrapping method calculates 10,000 times LIs using different thresholds ranging from zero until the maximum *t*-value for a specific contrast condition. For each threshold a cut-off mean value is obtained from which a weighted mean (LI-wm) index value can then be calculated (Wilke and Lidzba, 2007). This yields a single value between -1 and 1 indicating right- or left-sided hemisphere dominance.

Results

Our first goal was to identify brain regions involved in the processing of conspecific vocalizations by the macaque brain. To this end, we collected functional MR images of two monkeys in a horizontal 3-T scanner while stimuli from three different sound categories were presented to the animals. Complex sounds are characterized by having a wide range of spectrotemporal features. While environmental sounds typically contain sharp temporal onsets, monkey vocalizations contain greater modulations in the spectral domain because of the harmonics contained in these sounds. Environmental sounds also carry abstract information

about the identity of objects, so a comparison between BOLD responses to monkey vocalizations and environmental sounds is useful in determining brain structures involved in higher-level processing. However, specific spectrotemporal differences exist between these two types of sounds. This can be seen, for instance, in the spectral modulation of monkey vocalizations at approximately 1.5–2 cycles/kHz, which is not present for other sound categories (**Figure 1D**). Thus, scrambled versions of monkey calls (SMC) were used to further control for the local spectrotemporal features in the vocalizations (see **Figure 1C** and Material and Methods). Comparison of average modulation spectra between categories showed that SMC were acoustically better matched to

MC than Env (correlation coefficient between the modulation spectra: SMC vs. MC: 0.92, Env vs. MC: 0.86; **Figure 1D**).

Overall, sound stimulation elicited significant BOLD responses compared to silent trials irrespective of auditory stimulus category [q (FDR) < 0.05, p < 10^{-3} , one-tailed t -test, t range: 2.3–10, cluster size > 10 voxels] in a broad network of brain regions, including subcortical auditory pathways, classical auditory areas of the superior temporal gyrus (STG), but also regions in parietal and prefrontal cortices (**Figure 2**). The clusters in **Figure 2A** highlight the main activation sites on the cortical surface of monkey M1. **Figure 2B** shows selected coronal slices for both animals (M1 and M2) showing activation

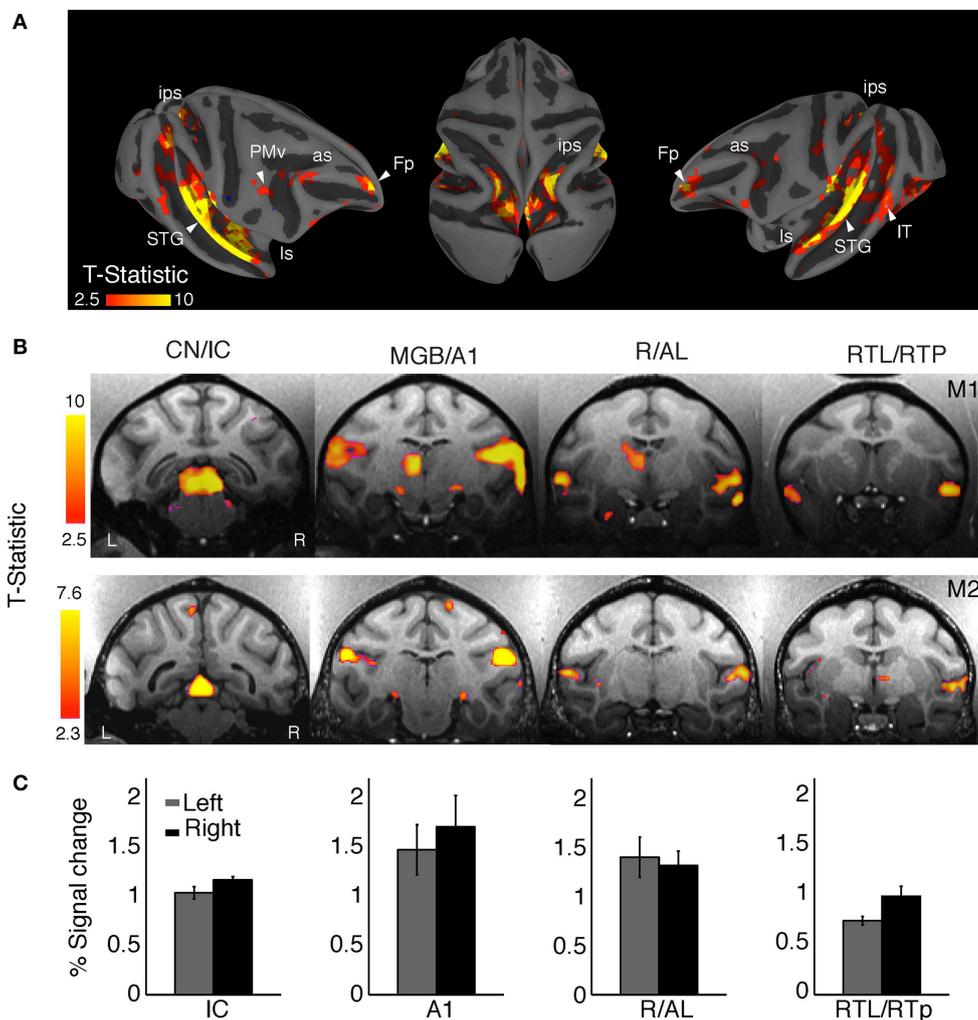


FIGURE 2 | Mapping auditory and auditory-related regions with complex sounds. (A) Representative cortical responses from monkey (M1) for all sound conditions combined (q FDR < 0.05, p < 10^{-2} ; cluster size > 10 voxels). The projection onto the semi-inflated surface preserves sulcal and gyral landmarks while allowing visualization inside the intraparietal sulcus (ips) and lateral sulcus (ls). Activation was observed along the auditory ventral stream in the superior temporal gyrus (STG), the superior temporal sulcus (STS), ventral intraparietal area (VIP), and the frontal pole (Fp). Activated dorsal-stream regions included the ips and

ventral premotor cortex (PMv). Some active clusters were also observed in the middle temporal area (MT) and the inferior temporal cortex (IT). **(B)** Activation was robust across regions in the ascending auditory pathway of the two monkeys: cochlear nuclei (CN), inferior colliculus (IC), medial geniculate nucleus (MGN), primary auditory cortex (A1), rostral area (R), anterolateral area (AL), lateral rostromtemporal area (RTL), and the rostromtemporal pole region (RTP). **(C)** The average BOLD response for the main auditory activation showed a right-hemisphere bias in both animals (M1, weighted mean = -0.33, M2, weighted mean = -0.66).

in the ascending auditory pathway. These regions include the cochlear nucleus (CN), the inferior colliculus (IC), the medial geniculate nucleus (MGN), the primary auditory cortex (A1), and areas in the anterior superior temporal cortex, including the rostral (R) and anterolateral (AL) areas, the rostrottemporal area (RTL), and the rostrottemporal pole (RTp) region.

Activation clusters (averaged across animals and hemispheres) taken from a normalized number of voxels (i.e., equal number of left and right voxels) were found in: IC [$N = 84$ voxels, peak coordinate = (4, -1, 12)]; A1 [$N = 198$ voxels, peak coordinate = (22, 6, 24)]; R/AL [$N = 131$ voxels, peak coordinate = (24, 17, 12)]; and RTL/RTp [$N = 165$ voxels, peak coordinate = (23, 22, 8)].

For both animals we observed a larger amplitude and spatial extent of the BOLD response in the right hemisphere as compared to the left hemisphere (Figure 2B). Activation (percent signal change) in selected clusters for each hemisphere is shown in Figure 2C. We compared the activation between the two hemispheres by calculating a laterality index (LI), with a positive index indicating a left-hemisphere bias and a negative index indicating a right-hemisphere bias. Given the fact that LIs show a threshold dependency (Nagata et al., 2001), we measured LI curves to provide a more comprehensive estimate over a whole range of thresholds (Wilke and Lidzba, 2007). Using this adaptive thresholding approach we found a right-hemisphere bias in the LI curves for general auditory activation (all sounds vs. baseline) in both monkeys (M1, weighted mean = -0.33; M2, weighted mean = -0.66). For higher thresholds, the activation was clustered in primary auditory cortex (A1) of the right hemisphere in each animal.

Vocalizations are complex naturalistic stimuli that contain behaviorally relevant information. In order to investigate if the

auditory system contained representations that are sensitive to this sound category vs. other types of behaviorally relevant complex sounds, we contrasted monkey calls against environmental sounds (see Material and Methods). Environmental sounds also carry abstract information about object identity in their spectrotemporal patterns. We, therefore, also looked for areas showing elevated response to these sounds relative to monkey vocalizations. When correcting for multiple comparisons [q (FDR) < 0.05], no differences were observed for the contrast of MC vs. Env. However, at uncorrected thresholds, we found significantly higher activations by MC as compared to Env in both monkeys across regions in temporal, parietal and prefrontal cortices (M1, $p < 10^{-3}$ uncorrected, t -value range: -4.2 to 6.1, cluster size > 5 voxels; M2, $p < 10^{-2}$ uncorrected, t -value range: -3.6 to 5.9, cluster size > 5 voxels) (Figure 3A). Specifically, activations sensitive to MC were found in the anterior STG region, including areas AL and RTp of the rostral belt/parabelt, and further along the auditory ventral stream in ventrolateral prefrontal cortex (vlPFC). In addition, we observed activation patches in the inferior parietal lobule (areas PF/PFG) of the right parietal cortex, and bilaterally inside the inferior branch of the arcuate sulcus, possibly corresponding to Brodmann's area (BA) 44, and posterior to the arcuate sulcus, in a region that is part of ventral premotor cortex (PMv). In addition, we found regions sensitive to environmental sounds (blue) along the superior temporal sulcus (STS) and inferotemporal (IT) cortex. To investigate hemispheric lateralization in the processing of vocalizations, we measured LI curves for this contrast (MC > Env), finding a slight right hemispheric bias in monkey M1 (weighted mean = -0.19) and a moderate right-hemisphere bias in monkey M2 (weighted mean = -0.42).

In order to determine whether spectrotemporal features alone could have driven the activation in these areas, we further

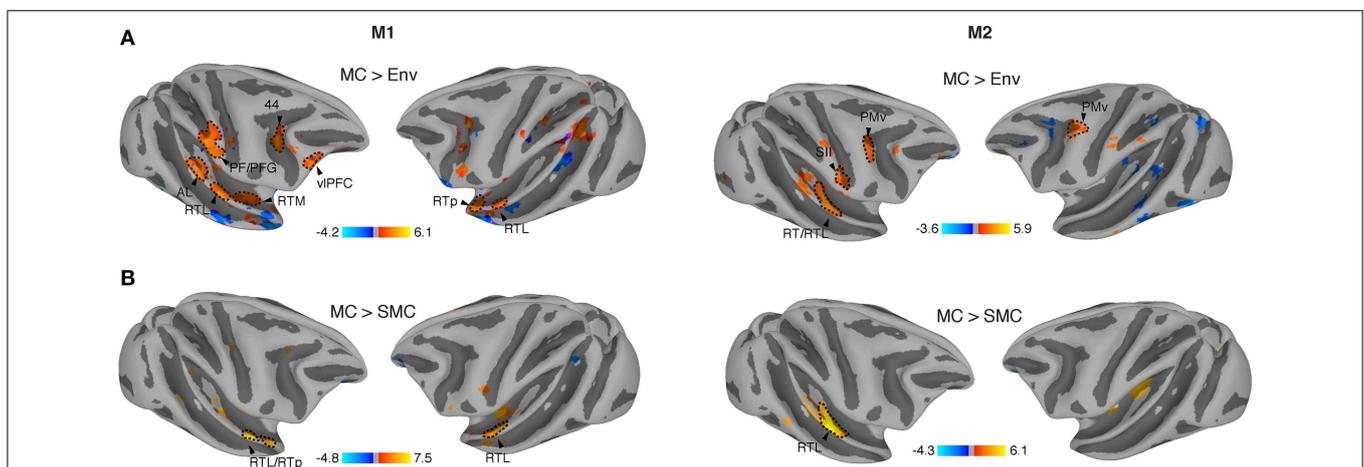


FIGURE 3 | Regions specifically activated by monkey vocalizations. (A) Vocalization-sensitive regions obtained from comparison between the effects of monkey calls and environmental sounds. All activation maps were displayed on a semi-flattened surface of the macaque monkey template. Active regions were found in the anterolateral area (AL), lateral rostrottemporal area (RTL), rostrottemporal pole (RTp), secondary somatosensory (SII)

cortex, ventral premotor cortex (PMv), ventrolateral prefrontal cortex (VIPFC), and inferior parietal areas (PF and PFG). (B) Regions significantly more activated by monkey vocalizations than by scrambled monkey vocalizations include areas in the anterior STG, RTL/RTp. Red/orange: significantly higher activation by MC than by control sounds (SMC or Env); blue: significantly higher activation by SMC or Env than by MC.

contrasted monkey calls (MC) with scrambled monkey calls (SMC). The results showed similar patterns of MC activation in both monkeys in the RTL region of the aSTG (M1; $p < 10^{-3}$ uncorrected, t -value range > -4.8 to 7.5 , cluster size > 5 voxels and for M2, $p < 10^{-2}$ uncorrected, t -value range > -4.3 to 6.1 , cluster size > 5 voxels) in both monkeys specifically in the RTL region of the aSTG (**Figure 3B**). In monkey M2, a second region, the middle medial belt (MB), was also more strongly activated by monkey vocalizations than by their scrambled counterparts. The weighted-mean lateralization index (LI) for this contrast (MC $>$ SMC) also showed higher values toward the right hemisphere (M1: weighted mean = -0.34 ; M2: weighted mean = -0.44). A summary is shown in **Table 1**.

Some differences in the patterns of activity were observed across the two animals. These differences might be explained either by variability across subjects or by differences in attentional state: M1 was significantly engaged in completing the task ($>90\%$ success), whereas M2 was scanned passively while holding fixation. To compensate for this variability, we calculated the minimum t -statistic ($p < 0.01$ uncorrected) across contrasts in each monkey (a conjunction test) and across monkeys in each contrast (**Figure 4**). Conjunction across contrasts (MC $>$ Env and MC $>$ SMC) and monkeys (M1 and M2) found a single area in the right hemisphere to be specifically involved across both conjunction analyses, area RTL/RTp (peak coordinate: 24, 17, 12).

Discussion

Species-specific vocalizations in non-human primates (“monkey calls”) convey important information about affective/emotional states as well as the recognition of objects and individuals (Ghazanfar and Hauser, 1999). We used whole-brain functional magnetic resonance imaging (fMRI) in awake behaving monkeys to examine auditory responses to stimuli from three different sound categories: (a) multiple types of conspecific monkey calls, (b) environmental sounds, and (c) scrambled versions of the same monkey calls largely preserving their local spectrotemporal features.

For all three sound categories combined we found robust BOLD responses along various regions in the ascending auditory pathways (CN, IC, MGB, and A1, **Figures 2A,B**). These results,

using a 3-T scanner without contrast agent, corroborate previous fMRI findings obtained on a 1.5-T magnet with the contrast agent MION, showing activation by complex sounds along the auditory pathway (Joly et al., 2012a). The results further attest to the fact that complex sounds are highly effective for mapping sub-cortical and cortical auditory structures (Rauschecker et al., 1995; Rauschecker, 1998b; Poremba et al., 2003). Furthermore, our results confirm the general trend of a slight right-hemisphere bias (**Table 1**) in the processing of complex sounds in the macaque auditory cortex, as measured with fMRI (Petkov et al., 2008; Joly et al., 2012a). Similar results have been found in humans for non-speech voice sounds (Belin et al., 2000).

When we compared activations produced by monkey vocalizations vs. the other two sound categories using a conjunction analysis, we found consistent activations in regions along the anterior STG, in particular in areas AL, RTL and RTp, in both animals (**Figure 4**). Our results extend previous findings of increased sensitivity to monkey vocalizations in anterior STG regions (Poremba et al., 2003; Petkov et al., 2008; Kikuchi et al., 2010; Joly et al., 2012a,b; Fukushima et al., 2014) by using control stimuli (SMC) that retained the low-level acoustic information of macaque vocalizations and whose acoustic structure was better matched to the vocalizations than the acoustic structure of other complex sounds (**Figure 1D**). Single-unit studies of the R/AL region have also found increased selectivity either to monkey calls, or to sound categories including vocalizations (Tian et al., 2001; Kuśmierk et al., 2012), consistent with the present results (**Figures 3, 4**).

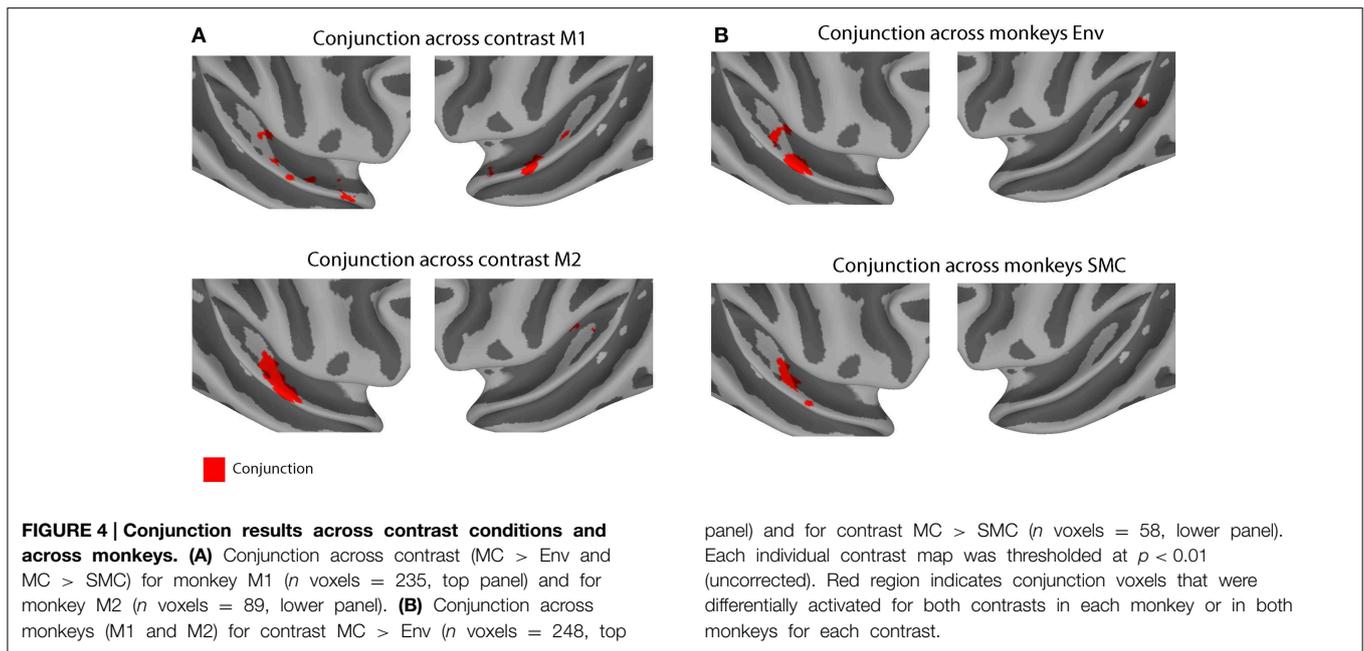
Thus, the cortical representation of vocalizations involves an auditory ventral pathway, consisting of a chain of interconnected regions in anterior STG and vlPFC that extract abstract information for the recognition and categorization of vocalizations (Rauschecker, 2012). The rostral belt, parabelt and aSTG send afferent projections into ventrolateral, polar, orbital, and medial regions of the prefrontal cortex (PFC) (Jones and Powell, 1970; Hackett et al., 1999; Romanski et al., 1999; Cavada et al., 2000; Kaas and Hackett, 2000; Hackett, 2011; Yeterian et al., 2012), and together these regions form the ventral cortical stream in audition. Vocalization-sensitive neurons are found along with face-sensitive neurons in the vlPFC (Romanski et al., 2005), allowing these regions to integrate vocalizations with the corresponding facial gestures (Romanski and Goldman-Rakic, 2002; Cohen et al., 2007; Diehl and Romanski, 2014). The PFC is involved in higher-level integrative processes for the cognitive control of vocalizations as well as in the interpretation of semantic content in vocalizations (Romanski and Averbeck, 2009). The activation patterns observed in PFC (**Figure 3A**) could represent categorical or affective information reflected in the vocalizations. Further imaging studies and multivariate analyses comparing multiple vocalization types might elucidate the differential contribution of each subregion of the PFC.

Our stimuli also activated higher-level visual areas, such as the middle temporal (MT) and inferior temporal areas (IT). These areas are known to be involved in the processing of visual motion (Maunsell and Van Essen, 1983) and in object perception (including faces), respectively (Tsao et al., 2006; Ku et al., 2011). Their activation by purely auditory stimuli raises interesting questions

TABLE 1 | LI-weighted-mean values for the overall sound activation and for each contrast condition.

	All $>$ baseline	MC $>$ Env	MC $>$ SMC
M1	-0.33	-0.19	-0.34
M2	-0.66	-0.42	-0.44

Mean lateralization index values (LI-wm) are shown that were obtained from LI curves measured as a function of the statistical threshold (t -value) for the overall auditory activation (all sounds vs. baseline), for the contrast between monkey calls and environmental sounds (MC $>$ Env) and for the contrast between monkey calls and scrambled monkey calls (MC $>$ SMC). A positive index indicates a left-hemisphere bias, while a negative index indicates a right-hemisphere bias. LI-wm values are shown separately for monkeys M1 and M2.



regarding their possible role in the multisensory processing of dynamic audio-visual stimuli, such as facial expressions that naturally occur in conjunction with vocalizations and/or motion of the face (Furl et al., 2012; Polosecki et al., 2013; Perrodin et al., 2014). However, to answer these questions more definitively, further imaging experiments utilizing dynamic audio-visual stimuli would be necessary. Such studies could enlighten us on how auditory information combines with visual information in both the ventral and dorsal pathways building multimodal representations from dynamic facial expressions combined with vocalizations (Ghazanfar and Logothetis, 2003).

When we contrasted monkey calls to environmental sounds, we also found differential activation in regions PF/PFG (area 7b) (Pandya and Seltzer, 1982; Rozzi et al., 2006) of the inferior parietal lobule (IPL), in addition to the well-known regions in the STG sensitive to monkey vocalizations. Parietal regions inside the intraparietal sulcus (IPS) have been known to receive auditory projections (Lewis and Van Essen, 2000) and to contain neurons that respond to auditory and multimodal stimuli (Stricanne et al., 1996; Bushara et al., 1999; Grunewald et al., 1999; Cohen and Andersen, 2000; Cohen, 2009), but the role of these regions has traditionally been assumed to lie in spatial processing and control of eye movements.

Similarly, we found an engagement of the ventral premotor cortex (PMv) in the processing of monkey vocalizations (Figure 3A). This region has previously been thought to be involved in the processing of the location (but not quality) of nearby sounds (Graziano et al., 1999). Surprisingly, when we compared the effects of vocalizations (MC) against vocalizations that were scrambled in both the spectral and temporal domains (SMC), we did not observe greater activation in parietal or prefrontal areas for MC, suggesting that the scrambled versions of the MC evoked the same amount of activity in these regions. Similar results were obtained by Joly et al. (2012b) with

temporally scrambled vocalizations activating large regions of premotor and parietal cortices. Ventral premotor cortex (PMv) has also been implicated in the initiation of vocalizations in the macaque monkey (Hage and Nieder, 2013). It appears possible, therefore, that the same neurons are the source of an efference copy signal (Kauramäki et al., 2010), which is responsible for the suppression of auditory cortex during self-initiated vocalizations (Eliades and Wang, 2003). More generally, they could be part of an audio-motor network connecting perception and production of sounds (Rauschecker and Scott, 2009; Rauschecker, 2011).

Author contributions

MO co-designed the study, trained the animals, programmed stimulus presentation, acquired part of the data, conducted most analyses, and co-wrote the manuscript. PK programmed the behavioral task and participated in writing the manuscript. DA trained the animals and acquired part of the data. ID generated the scrambled stimuli and acquired part of the data. FA contributed with data analyses and participated in writing the manuscript. GK, interpreted data and participated in writing the manuscript. MS, IJ, and JR co-designed the study and participated in writing the manuscript.

Acknowledgments

Special thanks to Josie Cui for animal care and assistance with the experiments, and John VanMeter for fMRI data optimization. This work was supported by grants from the National Institutes of Health (R01-DC03489, R01-NS052494, and R56-NS052494 to JR), a PIRE Grant from the National Science Foundation (OISE-0730255 to JR), and a FiDiPro award from the Academy of Finland (JR).

References

- Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., et al. (2006). Task-modulated “what” and “where” pathways in human auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14608–14613. doi: 10.1073/pnas.0510480103
- Arnott, S. R., Binns, M. A., Grady, C. L., and Alain, C. (2004). Assessing the auditory dual-pathway model in humans. *Neuroimage* 22, 401–408. doi: 10.1016/j.neuroimage.2004.01.014
- Barbas, H. (1993). Organization of cortical afferent input to orbitofrontal areas in the rhesus monkey. *Neuroscience* 56, 841–864. doi: 10.1016/0306-4522(93)90132-Y
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309–312. doi: 10.1038/35002078
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–528. doi: 10.1093/cercor/10.5.512
- Bornkessel-Schlesewsky, I., Schlewsky, M., Small, S. L., and Rauschecker, J. P. (2015). Neurobiological roots of language in primate audition: common computational properties. *Trends Cogn. Sci.* 19, 142–150. doi: 10.1016/j.tics.2014.12.008
- Brown, J. C. (1991). Calculation of a constant Q spectral transform. *J. Acoust. Soc. Am.* 89, 425. doi: 10.1121/1.400476
- Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M.-J., Rauschecker, J. P., and Hallett, M. (1999). Evidence for modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nat. Neurosci.* 2, 759–766. doi: 10.1038/11239
- Cavada, C., Compañy, T., Tejedor, J., Cruz-Rizzolo, R. J., and Reinoso-Suárez, F. (2000). The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cereb. Cortex* 10, 220–242. doi: 10.1093/cercor/10.3.220
- Cheney, D. L., and Seyfarth, R. M. (1990). *How Monkeys see the World*. Chicago: University of Chicago Press.
- Cohen, Y. E. (2009). Multimodal activity in the parietal cortex. *Hear. Res.* 258, 100–105. doi: 10.1016/j.heares.2009.01.011
- Cohen, Y. E., and Andersen, R. A. (2000). Reaches to sounds encoded in an eye-centered reference frame. *Neuron* 27, 647–652. doi: 10.1016/S0896-6273(00)00073-8
- Cohen, Y. E., Theunissen, F., Russ, B. E., and Gill, P. (2007). Acoustic features of rhesus vocalizations and their representation in the ventrolateral prefrontal cortex. *J. Neurophysiol.* 97, 1470–1484. doi: 10.1152/jn.00769.2006
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173. doi: 10.1006/cbmr.1996.0014
- Dale, A. M., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9, 179–194. doi: 10.1006/nimg.1998.0395
- DeWitt, I., and Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proc. Natl. Acad. Sci. U.S.A.* 109, E505–E514. doi: 10.1073/pnas.1113427109
- Diehl, M. M., and Romanski, L. M. (2014). Responses of prefrontal multisensory neurons to mismatching faces and vocalizations. *J. Neurosci.* 34, 11233–11243. doi: 10.1523/JNEUROSCI.5168-13.2014
- Eliades, S. J., and Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J. Neurophysiol.* 89, 2194–2207. doi: 10.1152/jn.00627.2002
- Fischl, B., Sereno, M. I., and Dale, A. M. (1999). Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9, 195–207. doi: 10.1006/nimg.1998.0396
- Frey, S., Campbell, J. S., Pike, G. B., and Petrides, M. (2008). Dissociating the human language pathways with high angular resolution diffusion fiber tractography. *J. Neurosci.* 5, 11435–11444. doi: 10.1523/JNEUROSCI.2388-08.2008
- Frey, S., Mackey, S., and Petrides, M. (2014). Cortico-cortical connections of areas 44 and 45B in the macaque monkey. *Brain Lang.* 131, 36–55. doi: 10.1016/j.bandl.2013.05.005
- Fukushima, M., Saunders, R. C., Leopold, D. A., Mishkin, M., and Averbeck, B. B. (2014). Differential coding of conspecific vocalizations in the ventral auditory cortical stream. *J. Neurosci.* 34, 4665–4676. doi: 10.1523/JNEUROSCI.3969-13.2014
- Furl, N., Hadj-Bouziane, F., Liu, N., Averbeck, B. B., and Ungerleider, L. G. (2012). Dynamic and static facial expressions decoded from motion-sensitive areas in the macaque monkey. *J. Neurosci.* 32, 15952–15962. doi: 10.1523/JNEUROSCI.1992-12.2012
- Ghazanfar, A. A., and Logothetis, N. K. (2003). Neuroperception: facial expressions linked to monkey calls. *Nature* 423, 937–938. doi: 10.1038/423937a
- Ghazanfar, A., and Hauser, M. (1999). The neuroethology of primate vocal communication: substrates for the evolution of speech. *Trends Cogn. Sci.* 3, 377–384. doi: 10.1016/S1364-6613(99)01379-0
- Gil-da-Costa, R., Braun, A., Lopes, M., Hauser, M. D., Carson, R. E., Herscovitch, P., et al. (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. doi: 10.1073/pnas.0408077101
- Graziano, M. S., Reiss, L. A., and Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature* 397, 428–430. doi: 10.1038/17115
- Grunewald, A., Linden, J. F., and Andersen, R. A. (1999). Responses to auditory stimuli in macaque lateral intraparietal area. I. Effects of training. *J. Neurophysiol.* 82, 330–342.
- Hackett, T. A. (2011). Information flow in the auditory cortical network. *Hear. Res.* 271, 133–146. doi: 10.1016/j.heares.2010.01.011
- Hackett, T. A., Stepniewska, I., and Kaas, J. H. (1999). Prefrontal connections of the parabelt auditory cortex in macaque monkeys. *Brain Res.* 817, 45–58. doi: 10.1016/S0006-8993(98)01182-2
- Hage, S. R., and Nieder, A. (2013). Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nat. Commun.* 4:2409. doi: 10.1038/ncomms3409
- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., et al. (1999). “Sparse” temporal sampling in auditory fMRI. *Hum. Brain Mapp.* 7, 213–223.
- Jackson, L. L., Heffner, R. S., and Heffner, H. E. (1999). Free-field audiogram of the Japanese macaque (*Macaca fuscata*). *J. Acoust. Soc. Am.* 106, 3017–3023. doi: 10.1121/1.428121
- Joly, O., Pallier, C., Ramus, F., Pressnitzer, D., Vanduffel, W., and Orban, G. A. (2012b). Processing of vocalizations in humans and monkeys: a comparative fMRI study. *Neuroimage* 62, 1376–1389. doi: 10.1016/j.neuroimage.2012.05.070
- Joly, O., Ramus, F., Pressnitzer, D., Vanduffel, W., and Orban, G. A. (2012a). Interhemispheric differences in auditory processing revealed by fMRI in awake rhesus monkeys. *Cereb. Cortex* 22, 838–853. doi: 10.1093/cercor/bhr150
- Jones, E. G., and Powell, T. P. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain* 93, 793–820. doi: 10.1093/brain/93.4.793
- Kaas, J. H., and Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11793–11799. doi: 10.1073/pnas.97.22.11793
- Kauramäki, J., Jääskeläinen, I. P., Hari, R., Möttönen, R., Rauschecker, J. P., and Sams, M. (2010). Transient adaptation of auditory cortex organization by lipreading and own speech production. *J. Neurosci.* 30, 1314–1321. doi: 10.1523/JNEUROSCI.1950-09.2010
- Kikuchi, Y., Horwitz, B., and Mishkin, M. (2010). Hierarchical auditory processing directed rostrally along the monkey’s supratemporal plane. *J. Neurosci.* 30, 13021–13030. doi: 10.1523/JNEUROSCI.2267-10.2010
- Kikuchi, Y., Horwitz, B., Mishkin, M., and Rauschecker, J. P. (2014). Processing of harmonics in the lateral belt of macaque auditory cortex. *Front. Neurosci.* 8:204. doi: 10.3389/fnins.2014.00204
- Ku, S.-P., Tolia, A. S., Logothetis, N. K., and Goense, J. (2011). fMRI of the face-processing network in the ventral temporal lobe of awake and anesthetized macaques. *Neuron* 70, 352–362. doi: 10.1016/j.neuron.2011.02.048
- Kuśmierek, P., Ortiz, M., and Rauschecker, J. P. (2012). Sound-identity processing in early areas of the auditory ventral stream in the macaque. *J. Neurophysiol.* 107, 1123–1141. doi: 10.1152/jn.00793.2011
- Kuśmierek, P., and Rauschecker, J. P. (2009). Functional specialization of medial auditory belt cortex in the alert rhesus monkey. *J. Neurophysiol.* 102, 1606–1622. doi: 10.1152/jn.00167.2009

- Kuśmierk, P., and Rauschecker, J. P. (2014). Selectivity for space and time in early areas of the auditory dorsal stream in the rhesus monkey. *J. Neurophysiol.* 111, 1671–1685. doi: 10.1152/jn.00436.2013
- Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., and Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *J. Neurosci.* 29, 2477–2485. doi: 10.1523/JNEUROSCI.4921-08.2009
- Lewis, J. W., and Van Essen, D. C. (2000). Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *J. Comp. Neurol.* 428, 112–137. doi: 10.1002/1096-9861(20001204)428:1<112::AID-CNE8>3.0.CO;2-9
- Maeder, P. P., Meuli, R. A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J. P., et al. (2001). Distinct pathways involved in sound recognition and localization: a human fMRI study. *Neuroimage* 14, 802–816. doi: 10.1006/nimg.2001.0888
- Maunsell, J. H., and Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *J. Neurophysiol.* 49, 1127–1147.
- McLaren, D. G., Kosmatka, K. J., Oakes, T. R., Kroenke, C. D., Kohama, S. G., Matochik, J. A., et al. (2009). A population-average MRI-based atlas collection of the rhesus macaque. *Neuroimage* 45, 52–59. doi: 10.1016/j.neuroimage.2008.10.058
- Mishkin, M., Ungerleider, L. G., and Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends Neurosci.* 6, 414–417. doi: 10.1016/0166-2236(83)90190-X
- Nagata, S. I., Uchimura, K., Hirakawa, W., and Kuratsu, J. I. (2001). Method for quantitatively evaluating the lateralization of linguistic function using functional MR imaging. *AJNR. Am. J. Neuroradiol.* 22, 985–991. Available online at: <http://www.ajnr.org/content/22/5/985.long>
- Pandya, D. N., and Seltzer, B. (1982). Intrinsic connections and architectonics of posterior parietal cortex in the rhesus monkey. *J. Comp. Neurol.* 204, 196–210. doi: 10.1002/cne.902040208
- Perrodin, C., Kayser, C., Logothetis, N. K., and Petkov, C. I. (2011). Voice cells in the primate temporal lobe. *Curr. Biol.* 21, 1408–1415. doi: 10.1016/j.cub.2011.07.028
- Perrodin, C., Kayser, C., Logothetis, N. K., and Petkov, C. I. (2014). Auditory and visual modulation of temporal lobe neurons in voice-sensitive and association cortices. *J. Neurosci.* 34, 2524–2537. doi: 10.1523/JNEUROSCI.2805-13.2014
- Petkov, C. I., Kayser, C., Stuedel, T., Whittingstall, K., Augath, M., and Logothetis, N. K. (2008). A voice region in the monkey brain. *Nat. Neurosci.* 11, 367–374. doi: 10.1038/nn2043
- Petkov, C. I., Kikuchi, Y., Milne, A., Mishkin, M., Rauschecker, J. P., and Logothetis, N. K. (2015). Different forms of effective connectivity in primate frontotemporal pathways. *Nat. Commun.* 6:6000. doi: 10.1038/ncomms7000
- Petrides, M., and Pandya, D. N. (2009). Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biol.* 7:e1000170. doi: 10.1371/journal.pbio.1000170
- Polosecki, P., Moeller, S., Schweers, N., Romanski, L. M., Tsao, D. Y., and Freiwald, W. A. (2013). Faces in motion: selectivity of macaque and human face processing areas for dynamic stimuli. *J. Neurosci.* 33, 11768–11773. doi: 10.1523/JNEUROSCI.5402-11.2013
- Poremba, A., Saunders, R. C., Crane, A. M., Cook, M., Sokoloff, L., and Mishkin, M. (2003). Functional mapping of the primate auditory system. *Science* 299, 568–572. doi: 10.1126/science.1078900
- Rauschecker, J. P. (1998a). Parallel processing in the auditory cortex of primates. *Audiol. Neurootol.* 3, 86–103. doi: 10.1159/000013784
- Rauschecker, J. P. (1998b). Cortical processing of complex sounds. *Curr. Opin. Neurobiol.* 8, 516–521. doi: 10.1016/S0959-4388(98)80040-8
- Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hear. Res.* 271, 16–25. doi: 10.1016/j.heares.2010.09.001
- Rauschecker, J. P. (2012). Ventral and dorsal streams in the evolution of speech and language. *Front. Evol. Neurosci.* 4:7. doi: 10.3389/fnevo.2012.00007
- Rauschecker, J. P., and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724. doi: 10.1038/nn.2331
- Rauschecker, J. P., and 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. doi: 10.1073/pnas.97.22.11800
- Rauschecker, J. P., Tian, B., and Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268, 111–114. doi: 10.1126/science.7701330
- Romanski, L. M., and Averbeck, B. B. (2009). The primate cortical auditory system and neural representation of conspecific vocalizations. *Annu. Rev. Neurosci.* 32, 315–346. doi: 10.1146/annurev.neuro.051508.135431
- Romanski, L. M., Averbeck, B. B., and Diltz, M. (2005). Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. *J. Neurophysiol.* 93, 734–747. doi: 10.1152/jn.00675.2004
- Romanski, L. M., and Goldman-Rakic, P. S. (2002). An auditory domain in primate prefrontal cortex. *Nat. Neurosci.* 5, 15–16. doi: 10.1038/nn781
- Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., and Rauschecker, J. P. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* 2, 1131–1136. doi: 10.1038/16056
- Rozzi, S., Calzavara, R., Belmalih, A., Borra, E., Gregoriou, G. G., Matelli, M., et al. (2006). Cortical connections of the inferior parietal cortical convexity of the macaque monkey. *Cereb. Cortex* 16, 1389–1417. doi: 10.1093/cercor/bhj076
- Russ, B. E., Ackelson, A. L., Baker, A. E., and Cohen, Y. E. (2008). Coding of auditory-stimulus identity in the auditory non-spatial processing stream. *J. Neurophysiol.* 99, 87–95. doi: 10.1152/jn.01069.2007
- Saleem, K. S., and Logothetis, N. K. (2012). *A Combined MRI and Histology Atlas of the Rhesus Monkey Brain in Stereotaxic Coordinates, 2nd Edn., Horizontal, Coronal and Sagittal Series.* San Diego, CA: Elsevier/Academic Press.
- Schörkhuber, C., and Klapuri, A. (2010). “Constant-Q transform toolbox for music processing,” in *7th Sound and Music Computing Conference* (Barcelona). Available online at: <http://www.iem.at/~schoerhuber/cqt2010/>
- Singh, N. C., and Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *J. Acoust. Soc. Am.* 114, 3394–3411. doi: 10.1121/1.1624067
- Stricanne, B., Andersen, R. A., and Mazzoni, P. (1996). Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *J. Neurophysiol.* 76, 2071–2076.
- Tian, B., Reser, D., Durham, A., Kustov, A., and Rauschecker, J. P. (2001). Functional specialization in rhesus monkey auditory cortex. *Science* 292, 290–293. doi: 10.1126/science.1058911
- Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., and Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science* 311, 670–674. doi: 10.1126/science.1119983
- Wilke, M., and Lidzba, K. (2007). LI-tool: a new toolbox to assess lateralization in functional MR-data. *J. Neurosci. Methods* 163, 128–136. doi: 10.1016/j.jneumeth.2007.01.026
- Wilson, S. M., Saygin, A. P., Sereno, M. I., and Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nat. Neurosci.* 7, 701–702. doi: 10.1038/nn1263
- Yeterian, E. H., Pandya, D. N., Tomaiuolo, F., and Petrides, M. (2012). The cortical connectivity of the prefrontal cortex in the monkey brain. *Cortex* 48, 58–81. doi: 10.1016/j.cortex.2011.03.004
- Yovel, G., and Belin, P. (2013). A unified coding strategy for processing faces and voices. *Trends Cogn. Sci.* 17, 263–271. doi: 10.1016/j.tics.2013.04.004

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Ortiz-Rios, Kuśmierk, DeWitt, Archakov, Azevedo, Sams, Jääskeläinen, Keliris and Rauschecker. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.