## **Supporting Information**

## Remedios et al. 10.1073/pnas.0909756106

## **SI Materials and Methods**

Behavioral and fMRI studies were performed with male rhesus macaques (*Macaca mulatta*) that are part of a colony housed at the Max-Planck Institute for Biological Cybernetics. All procedures were approved by the local authorities (Regierungspräsidium) and are in full compliance with the guidelines of the European Community (EUVD 86/609/EEC) for the care and use of laboratory animals.

Animals are socially housed in troops of 2–4 animals per enclosure, allowing for the establishment of natural hierarchical relationships. In addition to interactions within a troop, animals in one enclosure could observe animals in several other enclosures and communicate with each other by using visual and acoustic expressions. Strong gesturing, vocalizing, and territorial posturing between troops in adjacent enclosures was not uncommon. Enclosures are part of two separated facilities, allowing us to sort monkeys as either being familiar or unfamiliar to each other.

Acquisition of Sounds and Videos. Sounds from the animal colony were recorded while the monkeys were free within their enclosures by using a microphone (PC microphone; Telex) connected to a notebook computer and saved at 44,100 Hz (32 bit, wav files). The actual sound intensity was measured simultaneously using a condenser microphone (Brüel & Kjær 4188 and a 2238 Mediator sound level meter; Brüel & Kjær GmbH). For use in experiments, these sounds were normalized for intensity (total RMS power) and clipped to an adequate length by using Adobe Audition v2.0 (Adobe Systems).

For the behavioral and imaging experiments, four categories of sounds were used: vocalizations, drumming sounds, cage sounds, and other natural sounds. Drumming and cage sounds were especially recorded for this study. Conspecific vocalizations and other natural sounds were taken from previous studies (1, 2). Conspecific macaque vocalizations comprised coos, grunts, pant-threats, and screams. Cage sounds were sounds generated by items familiar to the animals, such sounds of a cage door, a primate-chair, or a partition being drawn. When measured, cage sounds were produced at 60–65 dB, whereas drumming sounds were much louder (81–88 dB). Other natural sounds comprised environmental and animal sounds, such as bubbling water, crickets, and vocalizations of birds and feline mammals.

Videos were recorded at 1,024 by 768 pixels at 25 frames per s by using a digital camera set on a tripod. Videos were clipped and arranged by using Adobe Premiere Pro v2.0 (Adobe Systems) and saved in Microsoft AVI format. Only clips in which the filmed animals were proximal to the cage front and with neutral expressions were considered. The relative sizes of the monkeys in the videos were maintained.

**Behavioral Tests.** Behavioral tests were based on preferential looking techniques, which are widely used to test perception in prelinguistic children and nonhuman primates (3–6). For testing, the subjects were seated in a primate chair placed in an anechoic and dark booth. Subjects were naive to the task and were not trained to perform in any respect. Importantly, the subjects were free to move their eyes and head to spontaneously orient toward stimuli of interest. Visual stimuli were presented by using a LCD monitor (117 cm; 46 in, Sharp Aquos LC-46, 60 cm in front of the subject), and sounds were presented by using two speakers set at head level, 50 cm to the left and to the right of the subject's head. The intensity of sound presentation was calibrated (see

Remedios et al. www.pnas.org/cgi/content/short/0909756106

below). All of the tests were performed in the dark, and the subject's natural behavior was recorded on digital video by using infrared light sources and cameras. Posthoc analysis was carried out in Matlab (MathWorks).

In the first experiment, we probed whether the animals preferentially orient toward sounds presented in the dark (Fig. 2A). Sounds were presented from either the left or right speaker, in a pseudorandom sequence and at random intervals. An equal number of sounds (n = 4) from each category was used, and the side of presentation was randomized across subjects. All sounds were normalized to the same intensity 75 dB r.m.s., and to eliminate spurious influences from potential background noises and to restrict behavioral reactions to salient sounds, a continuous 60-dB masking sound (music) was continuously played as background. This choice of background sound was deliberate, so as to mimic realistic conditions, because music is regularly played in the background in the animal facilities. The orienting behavior of each subject subsequent to a sound presentation was scored by an observer that was blind to the direction or kind of sound presented. An orienting response was noted if the animal turned the head or eyes toward the side of any speaker within 5 s from the sound presentation. As in previous studies, such orienting responses were unambiguous (3, 4). For each sound category, the percentage of sounds with orienting response toward the speaker presenting the sound was scored. Fig. 2A displays the distribution of this percentage across animals and sound categories.

In the second experiment, we probed the influence of drumming or control cage sounds on the animals' preferential looking behavior toward videos showing other conspecifics. Two videos were presented on the far left and right side of the screen (each covering  $\approx 20^{\circ}$  of visual angle), one showing a large, the other a small animal. The large animals were significantly older and larger than the tested subjects, whereas the small animals were younger and smaller than these. The sides of presentation were balanced across subjects. A binaural sequence of drumming or cage sounds (5-s intersound interval) was presented continuously for 1 min, where each sound was presented at its "natural" intensities, which had been measured in the colony (85 dB for drumming and 65 dB for cage sounds). Importantly, although some of the subjects were unfamiliar with both monkeys shown in the movies, other subjects were familiar with either of them. Familiarity here was defined as being housed either in adjacent enclosures, ensuring regular acoustic or visual interactions. The looking behavior of each subject was scored during the 1-min stimulus presentation by an observer blind to the experimental condition. For each trial, the total time (in seconds) spent viewing either movie (or away from both) was computed. Fig. 2B displays the distribution of this total viewing time across animals.

**Functional Imaging.** Functional imaging experiments were performed with anesthetized animals and on a high-field magnet (4.7 T, Biospec 47/40v; Bruker Medical) by using established protocols (1, 7–9). Anesthesia was maintained by using fastacting opiods (remifentanyl 0.5–2  $\mu$ g/kg per min), the animals were ventilated, and body temperature, blood oxygenation, and pressure and heart rate were continuously monitored and kept in the desired range. Signals were acquired by using whole-head volume coils, and MR-compatible headphones were positioned over both ears and sealed with sound attenuating foam. We acquired functional data from three animals (two imaged twice).

Image acquisition and acoustic stimulation protocols largely followed those used in previous studies (1, 7, 10). Functional data

were acquired with a multishot (two segments) gradient-recalled echo planar imaging sequence with typical parameters (TE, 20 ms; volume acquisition time TA, 1 s; effective TR, 10 s; flip angle, 60; spectral width, 100 kHz; on a grid of  $128 \times 128$  voxels; 2-mm slice thickness; field of view,  $96 \times 96$  mm, 9-12 slices). Anatomical images (T1-weighted) were acquired with a foursegment MDEFT sequence (TE, 4 ms; TR, 15 ms; flip angle, 20; spectral width, 75 kHz;  $384 \times 384$  voxels). The image slices were oriented parallel to the lateral sulcus to capture auditory cortex within the smallest number of slices. Anatomical and functional images were acquired on the same field of view, to facilitate superposition. Sounds were presented at an average intensity of 80 dB and were chosen from the same dataset as used in the behavioral experiments: drumming sounds, conspecific vocalizations, and other natural sounds (comprising animal vocalizations and environmental sounds). Different sound categories were presented in a pseudorandom order. During a stimulation interval, random sounds of the respective category were presented for a 40-s period (200-ms intersound interval), followed by 40-s baseline. Functional images were acquired by using a sparse imaging sequence every 10 s. This procedure leaves

- 1. Petkov CI, et al. (2008) A voice region in the monkey brain. Nat Neurosci 11:367-374.
- Kuhl PK, Meltzoff AN (1982) The bimodal perception of speech in infancy. Science 218:1138–1141.
- Ghazanfar AA, Logothetis NK (2003) Neuroperception: Facial expressions linked to monkey calls. Nature 423:937–938.
- Ghazanfar AA, Turesson HK, Maier JX, van Dinther R, Patterson RD, Logothetis NK (2007) Vocal-tract resonances as indexical cues in rhesus monkeys. *Curr Biol* 17:425–430.
- Patterson ML, Werker JF (1999) Matching phonetic information in lips and voice is robust in 4.5 month-old infants. *Infant Behav Dev* 22:237–247.
- McDermott J, Hauser MD (2007) Nonhuman primates prefer slow tempos but dislike music overall. *Cognition* 104:654–668.

sufficient time for acoustic stimulation in the absence of scanner noise.

Functional data were analyzed in Matlab (MathWorks). Time series were detrended, normalized to units of percentage signal change (vs. baseline), and two contrasts for voxels preferentially responding to vocalizations or drumming sounds constructed. The first contrast compared activations to conspecific vocalizations with those to other natural sounds, without constraints on the activation to drumming sounds. The second contrast compared activations to drumming sounds with those to other natural sounds, without constraints on the response to vocalizations. Voxels with significant effects in either contrast were identified by using spatial clustering (requiring eight active voxels in a  $4 \times 4$  neighborhood). The corresponding P values were computed by using a randomization procedure, by using voxels outside the brain, to ensure a false-discovery-rate of active voxels of <5% (9). Subsequent preference maps for vocalizations and drumming sounds were displayed superimposed on the anatomical MR images (Fig. S3). To determine those regions where preferences to conspecific vocalizations and drumming sounds overlap, we determined the overlap of both preference maps (as shown in Fig. 3 and Fig. S4).

- Kayser C, Petkov CI, Augath M, Logothetis NK (2007) Functional imaging reveals visual modulation of specific fields in auditory cortex. J Neurosci 27:1824–1835.
- Logothetis NK, Guggenberger H, Peled S, Pauls J (1999) Functional imaging of the monkey brain. Nat Neurosci 2:555–562.
- Tolias AS, et al. (2005) Mapping cortical activity elicited with electrical microstimulation using fMRI in the macaque. *Neuron* 48:901–911.
- Kayser C, Petkov CI, Lippert M, Logothetis NK (2005) Mechanisms for allocating auditory attention: An auditory saliency map. *Curr Biol* 15:1943–1947.
- 11. Petkov Cl, Kayser C, Augath M, Logothetis NK (2006) Functional imaging reveals numerous fields in the monkey auditory cortex. *PLOS Biol* 4:e215.



**Fig. S1.** (*A*) Examples of drumming sounds created by different individuals. The regular "beats" constituting drumming sounds are well visible, although their number differs between each of the sounds. (*B*) The spectral density at frequencies <10 kHz is well conserved across drumming samples (n = 13). (*C*) Examples of different macaque vocalizations. The spectrograms illustrate that vocalizations are acoustically distinct from drumming sounds. Note that pant-threats have an apparent beat-like structure, but on a much shorter time scale than drumming sounds and with variable and irregular interbeat intervals that depend on the individual vocalization. (*D*) Cage sounds produced by inanimate objects feature comparable abrupt acoustic features as drumming sounds but lack the periodic beat pattern.



**Fig. 52.** Drumming sounds are produced in different behavioral contexts. Most frequent were individual agonistic situations ("agonistic drumming"), in which one animal produced a directed drumming display toward another, either as a threatening display or as a display of power and dominance. In such situations, the drummer would often show threatening facial expressions (Fig. 1*A*), and drumming was often accompanied by threatening vocalizations elicited by the drummer. In the example, two members of a troop addressed an animal in an adjacent enclosure by using drumming sounds and pant-threat vocalizations. Drumming sounds were also produced in polyagonistic contexts ("polyagonistic drumming"), which involved several individuals from different enclosures producing displays of dominance and threat. Often such situations were elicited by one animal performing the display and were reciprocated by other animals from possibly distant enclosures. Sometimes two animals performing drumming sounds would not have direct eye contact and rely on the sound as a means to address the other. Vocalizations (threats) were also produced in situations. In the example, animals from several enclosures engaged in polyagonistic drumming sounds would not have direct eye contact and rely on the sound as a means to address the other. Vocalizations (threats) were also prequent in such situations of general excitement ("excited drumming"), such as shortly before feeding. Here, some animals performed drumming sounds, whereas others produced a chorus of vocalizations, ranging from multicontext calls and threats to food-associated calls. In the example, one animal spotted a caretaker with a bowl of fruits and started the excitement phase by a harmonic arch vocalization. Other animals continued with a chorus of vocalizations and drumming. The pie chart indicates the relative frequency of different contexts, extracted from n = 50 observations of drumming behavior on 7 days. Colored boxes indicate periods containing vocalizations, drumming, or



**Fig. S3.** (*A*) Regions responding preferentially to vocalizations and drumming sounds for one of the animals (the same a shown in Fig. 3). The left panel reveals clusters of voxels in the temporal lobe of both hemispheres responding preferentially to vocalizations compared to other natural sounds (shown in red). The right panel reveals clusters responding preferentially to drumming sounds compared to natural sounds. For direct comparison, the same drumming-preferring clusters are also shown in the left panel (as white outlines superimposed over the red vocalization-preferring regions). Clear overlap of both groups are visible in the caudal auditory cortex (blue arrows), the amygdaloid region (green arrows), and in the putamen (yellow arrow). The latter, however, was not consistent across animals. Note that image slices have been acquired parallel to the lateral sulcus, as shown in the schematic (*Bottom Right*). Statistical *P* values are color-coded and have been corrected for cluster size and false discovery rates. (*B*) Enlarged display of slices 3 and 4 from *A*.



В

N A N C

Caudal auditory cortex region - animal 3

Functional parcellation of auditory cortex



Tpt: Temporo-parietal region, Ts1: region temporalis superior 1, Ts2: region temporalis superior 2

**Fig. 54.** (*A*) Overlap of regions preferentially responding to vocalizations and drumming sounds compared to other natural sounds. As shown in **Fig. 53**, individual contrasts for vocalizations and drumming sounds were computed and the overlap of both determined. The above displays these overlap regions for animals 2 and 3, in analogy to Fig. 3. Regions falling in caudal auditory cortex are outlined in blue, and those in the amygdaloid region are outlined in green. Along with Fig. 3, this demonstrates that networks preferentially responding to vocalizations or drumming sounds consistently overlap in the caudal auditory cortex and the amygdala. Regions outlined in yellow show an overlap for individual animals, but did not consistently appear in all animals tested. (*B*) Functional localization of overlap regions. The left panel displays the caudal auditory region in animal 3 (as in *A*) together with a functional parcellation of auditory cortex (red outlines), obtained by using a described procedure (12). This provides a functional localization of individual fields in single subjects by mapping their tonotopic (sound frequency) organization. The core, belt, and parabelt fields have been color-coded in the right panel. The caudal regions showing overlapping preference to conspecific vocalizations and drumming sounds were located mostly in the caudal belt (fields CM and CL), reaching somewhat into primary field A1 and the temporoparietal region Tpt. Note that in animal 3 there is a second small cluster. Located more anterior in fields A1 and R. The localization of a vocalization preferring region in the caudal belt is in good agreement with a previous study (1).