Title: Level dependence of spatial processing in primate auditory cortex

Authors: Yi Zhou and Xiaoqin Wang

Affiliation: Laboratory of Auditory Neurophysiology, Department of Biomedical Engineering, School of Medicine, Johns Hopkins University, Baltimore, MD 21205

Running head: Space coding in auditory cortex

Contact information:
Yi Zhou, PhD
Laboratory of Auditory Neurophysiology
Department of Biomedical Engineering
School of Medicine, Johns Hopkins University
720 Rutland Avenue, Traylor 419
Baltimore, MD 21205-2195, U.S.A.
Tel (410) 614-4547 / Fax (410) 614-9599

Key words: space coding; marmoset; auditory cortex; single neuron; sound localization
Abstract

Sound localization in both humans and monkeys is tolerant to changes in sound level. The underlying neural mechanism, however, is not well understood. This study reports the level dependence of individual neurons' spatial receptive field (SRF) in the primary auditory cortex (A1) and adjacent caudal field in awake marmoset monkeys. We found that most neurons' excitatory SRF components were spatially confined in response to broadband noise stimuli delivered from the upper frontal sound field. Approximately half the recorded neurons exhibited little change in spatial tuning width over a ~20dB change in sound level, while the remaining neurons showed either expansion or contraction in their tuning widths. Increased sound levels did not alter the percent distribution of tuning width for neurons collected in either cortical field. The population averaged responses remained tuned between 30 and 80dB SPL for neuronal groups preferring contralateral, midline, and ipsilateral locations. We further investigated the spatial extent and level dependence of the suppressive component of SRFs using a pair of sequentially presented stimuli. Forward suppression was observed when the stimuli were delivered from “far” locations, distant to the excitatory center of an SRF. In contrast to spatially confined excitation, the strength of suppression typically increased with stimulus level at both excitatory center and far regions of an SRF. These findings indicate that although the spatial tuning of individual neurons varied with stimulus level, their ensemble responses were level tolerant. Widespread spatial suppression may play an important role in limiting the sizes of SRFs at high sound levels in the auditory cortex.
Introduction

In auditory perception, the accuracy of sound localization shows remarkable tolerance to changes in sound level (Sabin et al. 2005) with a general degradation towards threshold levels in humans (Su and Recanzone 2001; Vliegen and Van Opstal 2004) and monkeys (Recanzone and Beckerman 2004). Using anesthetized preparations, electrophysiological studies have amassed considerable evidence for the existence of location-sensitive neurons in the auditory cortex (Ahissar et al. 1992; Brugge et al. 1996; Eggermont and Mossop 1998; Imig et al. 1990; Middlebrooks and Pettigrew 1981; Middlebrooks et al. 1998; Mrsic-Flogel et al. 2005; Rajan et al. 1990; Reale et al. 2003). In these studies, the spatial tuning of the majority of cortical neurons is found to be sharpest at near-threshold level, progressively broadening at moderate to high sound levels (e.g., Brugge et al. 1996; Middlebrooks et al. 1998; Mrsic-Flogel et al. 2005; Xu et al. 1998). This trend conflicts with documented perceptual performance in sound localization. Theoretical studies have suggested the use of disparity information between two broadly tuned neural populations to encode azimuth (Stecker et al. 2005).

More recently, spatial selectivity has been reevaluated in the primary auditory cortex (A1) and secondary auditory cortex using awake and behaving preparations (King et al. 2007; Lee and Middlebrooks 2010; Mickey and Middlebrooks 2003; Recanzone 2000; Werner-Reiss and Groh 2008; Woods et al. 2006). In contrast to the general findings in anesthetized animals, spatial tuning width obtained in awake preparations does not systematically expand with increasing sound level (Mickey and Middlebrooks 2003; Woods et al. 2006). In fact it sharpens through the suppression of responses at less preferred locations during behavioral tasks (Lee and Middlebrooks...
These studies have suggested that inhibition plays a role in limiting the size of SRFs with increasing sound level. In principle, one may infer the presence of inhibition from a reduction in neural firing under proper conditions. However, inhibition is generally difficult to characterize in extracellular recordings due to low spontaneous firing rates, especially in the upper cortical layers of awake animals (typically <2 spike/sec in awake marmosets, Barbour and Wang 2003; Wang et al. 2005). The relationship between the strength of inhibition (or response suppression) and sound level across spatial locations has not been systematically examined in the auditory cortex of awake animals.

Another contrasting finding for awake animals is that the azimuthal tuning widths of single- and multi-unit activity could either expand or contract with an average 20dB increase in sound level in the auditory cortex of cats (Mickey and Middlebrooks 2003). Yet the overall distribution of tuning width appears to be unchanged between moderate and high sound levels in the auditory cortex of macaques (Woods et al. 2006). It should be noted that expansion and contraction in individual SRFs do not warrant a level-invariant distribution of tuning widths. For example, if broad SRFs become broader and narrow SRFs become narrower, the overall distribution of tuning widths may become bimodal at higher sound levels. Conversely, convergence may occur. The nature of level invariance may differ between responses of individual neurons and their population average, depending on the distribution of tuning widths across sound level. This distinction is important because tuning width directly influences the stimulus information carried by a population of neurons (Kang et al. 2004; Seung and Sompolinsky 1993; Zhang and Sejnowski 1999). To address this issue, systematic analyses are needed to clarify the relationship between the variability found in individual
neurons’ spatial tuning widths and the overall tuning properties of their ensemble responses in the auditory cortex of awake animals.

In recent years, the marmoset model has offered many insights into spectral and temporal aspects of sound processing in the auditory cortex of awake animals (e.g., Barbour and Wang 2003; Lu et al. 2001b). However, cortical processing of sound location information remains relatively unknown in the marmoset species. To improve the value of this nonhuman primate model, this study investigated neurons’ spatial tuning in the primary auditory cortex (A1) and the adjacent caudal field (CM/CL) of awake marmoset monkeys. Our experiments focused on the two aforementioned issues for spatial coding: (1) the level dependence of spatial responses of individual neurons and their population average; and (2) the level dependence of suppression in cortical SRFs. The results show that although the spatial tuning width of individual neurons could expand or contract with increasing sound level, the firing rate versus azimuth profiles of averaged population responses remained tuned for neuronal groups preferring contralateral, midline, and ipsilateral directions. We further examined the spatial extent of suppression in cortical SRFs using a pair of sequentially presented test-probe stimuli. Forward suppression was observed when the test stimuli were presented from either excitatory center or far regions of a SRF. In general, the strength of suppression increased with sound level, even at far locations showing much reduced excitatory activity. These results provide insights into the roles of suppression mechanisms in spatial processing in the auditory cortex.
Materials and Methods

Animal preparation and electrophysiological procedures

A chronic recording preparation (e.g., Lu et al. 2001a) was used to record single-neuron activity in the auditory cortex of awake common marmoset monkeys (*Callithrix jacchus*). Experimental procedures were approved by the Institutional Animal Care and Use Committee of the Johns Hopkins University following NIH guidelines.

All subjects were trained to sit in a custom-designed primate chair. After 2-4 weeks of behavioral adaptation, two stainless steel headposts were attached to the skull under sterile conditions with the animal deeply anesthetized by isoflurane (0.5-2.0%, mixed with 50% oxygen and 50% nitrous oxide). The headposts served to maintain a stable head orientation of the subject during electrophysiological recordings. To access the auditory cortex, small craniotomies (~1 mm in diameter) were made on the skull over the superior temporal gyrus to allow for the penetration of electrodes (tungsten electrodes, 2-5 MΩ impedance, A-M systems, Carlsborg, WA) via a hydraulic microdrive (Trent-Wells, Los Angeles, CA). Single-unit activity was sorted online using a template-based spike-sorting program (MSD, Alpha Omega Engineering) and analyzed using custom programs written in Matlab (Mathworks, Natick, MA).

Spike waveforms were continuously monitored throughout a recording session to ensure stability and isolation quality. The 25, 50, and 75 percentiles of the signal-to-noise ratio (SNR) were 16.56, 19.43, and 23.66 dB, respectively, for neurons reported in this study. SNR was defined as $20\log_{10}(|V_{\text{max}}|/\sigma_n)$ in decibels, where $V_{\text{max}}$ was the maximal deflection of a spike waveform and $\sigma_n$ was the SD of baseline activity. The median variability of the SNR of a neuron was 1.59 dB (SD) across stimulus sets tested.
Experimental setup and sound delivery

Experiments were conducted in a dimly lit, double-walled, sound-attenuated chamber (IAC-1024, Industrial Acoustics, 1.9 x 2.2 x 1.9 m). The internal walls and ceiling were lined with ~3 inch acoustic absorption foam (Sonex, Illbruck) and the recording table and speaker frame were covered with ~1 inch acoustic absorption foam to reduce acoustic reflections. Figure 1A illustrates the experimental arrangement. The subject sat in the primate chair centered in the room. Fifteen loudspeakers (FT28D, Dome Tweeter, Fostex Co) were mounted on a semicircular frame covering the upper-level frontal field at a distance of ~80 cm from the head of a subject. The position of a loudspeaker was specified in azimuth (AZ) and elevation (EL) angles in spherical coordinates. As illustrated in Fig. 1B, a set of seven loudspeakers were evenly positioned with 30° horizontal spacing (AZ ±90°, ±60°, ±30°, 0°) at EL 0° and EL 45°; one loudspeaker was positioned directly above the head of a subject (EL 90°). The loudspeaker directly in front of the animal was at AZ 0° and EL 0°. Positive AZ angles corresponded to speakers ipsilateral to the recording site. During experiments, the subject’s head remained fixed and eye position was not controlled. An infrared camera was used to monitor the general behavior of a subject throughout a recording session.

All stimuli were generated digitally at a sampling rate of 100 kHz, converted via a digital-to-analog interface (Tucker-Davis Technologies, TDT, DA4), attenuated (TDT, PA4), amplified (Crown Amplifier, D75A), multiplexed (TDT, PM1, 2x8 channel), and
delivered to one or two loudspeakers (i.e., single or two-source stimulation) via a 16-channel breakout box (custom built, connected to PM1). For single-source stimulation, one stimulus was generated and PM1 was set at the “mono” mode (1x16) to deliver the stimulus to a designated loudspeaker. For two-source stimulation, two stimuli were generated and operated by separate TDT modules (PA4) or ports (DA4, PM1). In this case, PM1 was set to the “stereo” mode (channel A and B, 2x8) to deliver two stimuli (one per channel) simultaneously to two designated loudspeakers. As shown in Fig. 1B, the fifteen loudspeakers were wired to PM1 with an interleaved arrangement (black: channel A; gray: channel B), so that the stimuli from the two channels both covered the contralateral, midline, and ipsilateral sound field.

The loudspeaker impulse responses were obtained by $2^{14}$-point Golay code stimulation (Zhou et al. 1992), using a free-field microphone (B&K Type 4191; ½-in) placed on the top of the primate chair pointing towards the loudspeaker under test. Electrophysiological equipment such as the head-holder and electrode manipulator were removed during recording. The maximum deviation in power spectrum between responses of individual loudspeakers and their group average was within $\pm 6$dB/Hz in the frequency range of 2-32 kHz except for the top speaker, whose maximum deviation was about $\pm 12$ dB/Hz due to reflections from the primate chair. Due to the finite frequency-response ranges of loudspeakers in our setup, acoustic stimuli were limited to the frequency range 2-32 kHz, with a few exceptions for testing frequency tuning of neurons whose maximal responses occurred below 2 kHz or above 32 kHz.

Characterization of single-neuron response properties
After a neuron was isolated, pure-tone and noise stimuli (100 ms in duration with 5-ms cosine rise and fall times) were played iteratively to characterize its spatial, frequency, and level selectivity. Spatial selectivity was examined using broadband, frozen Gaussian noise (FIR filtered, flat spectrum between 2-32 kHz; at least 10 dB above the threshold). Noise samples were randomly chosen from neuron to neuron. Frequency selectivity was examined using pure-tone stimuli (2~32 kHz in 10 steps/octave) played at moderate sound pressure levels (SPLs, 30-60 dB) from at least one driven speaker location. BF was defined as the pure-tone frequency evoking the maximal rate of a neuron across the range of SPLs and locations tested. The average discharge rate was calculated based on the number of spikes occurring 0-150 ms after the stimulus onset.

Pure-tone and noise intensities were both expressed in terms of the peak-to-peak equivalent dB SPL. The reference amplitude was set by a 1 kHz tone calibrated at ~90 dB SPL (re. 20μPa) with zero-dB peak attenuation. The spectrum level of noise was 40 dB/Hz at zero-dB peak attenuation. In experiments, the inter-stimulus-interval was at least 500 ms for randomly presented stimulus sequences. Ten repetitions were played for each stimulus.

**Characterization of spatial selectivity of a neuron**

Multiple metrics were used to characterize the spatial sensitivity of a neuron. Best speaker location corresponded to the location evoking the maximal firing rate. Modulation depth (MD) corresponded to the peak-normalized difference between the maximal and minimal firing rates, \((R_{\text{max}} - R_{\text{min}})/R_{\text{max}}\), collected across fifteen speaker
locations. Average firing rate $R_{avg}$ across fifteen locations was used to examine the overall excitability of a neuron.

To facilitate the graphical inspection of neural activity and to quantify the size of tuning area in a two-dimensional (2-D) space, we constructed SRFs by the methods of interpolation and extrapolation. A similar approach has been used in studies of cortical SRFs in ferret (Mrsic-Flogel et al. 2005). Here we interpolated average rates over an equal-area grid consisting of an ensemble of squares of 5° resolution (0.087x0.087 in radius²) that evenly tiled the spherical surface. Since speaker distance was not a parameter of interest in this study, SRFs were mapped onto the unit sphere. The total surface area of the frontal hemifield (0° to 90° in EL and -90° to 90° in AZ) was 3.417, close to $\pi$, a quarter of the surface area of a unit sphere. The difference indicates the quantization error of the grid. To ensure accurate estimations of tuning area, the grid map was further extended 30° laterally and below (-120° to 120° in AZ and -30° to 90° in EL). This allowed a proper estimate of tuning area driven by loudspeakers located at the borders of the speaker array (e.g., EL0°); otherwise, erroneous smaller tuning areas would be reported compared to those driven by loudspeakers located at the center of the array (e.g., AZ0° EL45°).

In a SRF, activity at a grid position with $\theta$ in EL and $\phi$ in AZ, denoted as $(\theta,\phi)$ in spherical coordinates, was calculated as the weighted sum of responses over all speaker locations: $r(\theta,\phi) = \sum_{i=1}^{N} A_{i} R(\theta_{i},\phi_{i})$, where $R(\theta_{i},\phi_{i})$ was the average rate to the $i$th speaker at $(\theta_{i},\phi_{i})$. $A_{i}$ was an exponential function $A_{i} = \exp(-\alpha^2 / 2\sigma^2)$ with $\sigma=20^\circ$, $\alpha$ was the angular distance (in degrees) between $(\theta,\phi)$ and $(\theta_{i},\phi_{i})$, and $N=15$. The SRFs
plotted the peak-normalized \( r(\theta, \phi) \) and were flattened into a 2-D map for display with the abscissa for azimuth and the ordinate for elevation.

The accuracy of the interpolation/extrapolation method was evaluated by examining how closely the centroid of a simulated SRF matched the position of a single driving speaker (zero responses were assumed at all other speaker locations). Using the equal-area grid, the error was 0° in EL and less than 5° in AZ for speakers at EL 0° and 45°, and the error was 0° in AZ and 10° in EL for the top speaker. Since only a quarter of the spherical space was sampled, we quantified the spatial tuning width of a neuron, best area (BA), as the ratio between the area of a SRF with activity ≥62.4% of the peak and the area of the frontal field covered by the speaker array. Using this threshold criterion, a spatially highly selective neuron responding only to one speaker would yield BA of 0.1, while a spatially non-selective neuron responding equally to all fifteen speakers would yield BA of 0.996.

**Methods of two-source stimulation for characterizing suppression in SRFs**

Previous studies have suggested that suppression at less preferred locations sharpens the spatial tuning of a cortical neuron in the awake condition (Lee and Middlebrooks 2010; Mickey and Middlebrooks 2003). However, due to the low spontaneous rate, the strength of suppression is difficult to reveal in extracellular studies. The level dependence of suppression remains uncharacterized in cortical SRFs.

In principle, one may infer the tuning of suppression to a test stimulus (S1) from reductions in neural responses to a probe stimulus (S2) using the difference formula \( R(S1+S2) - R(S2) \). This method has been widely used in studies of frequency tuning of
excitation and inhibition in the auditory system (i.e., “two-tone” paradigm, Sachs and Kiang 1968, Suga and Tsuzuki 1985, Sutter et al. 1999). However, it is problematic to use a pair of simultaneously presented S1 and S2 to investigate the region of suppression in SRFs because perceptual fusion may occur (i.e., “summing localization”, Leakey 1959; Snow 1954). More specially, merging sound waves at the ears originating from two different directions could change the binaural coherence of the composite signal (Blauert 1997). This process cannot be approximated as simply addition between the probe and test stimuli.

For this reason, we characterized the rate-level tuning of excitation and suppression at multiple spatial locations using a pair of lead-lag sounds (S1 and S2). The experimental protocol was similar to the “forward masking” paradigm for studying excitatory and inhibitory frequency tuning in the auditory cortex (e.g., Brosch and Schreiner 1997; Calford and Semple 1995), except that S1 and S2 were delivered from either same or different locations. The test stimulus S1 was broadband, frozen Gaussian noise (band-limited between 2-32 kHz) played at SPLs ranging from -10 to 80dB in 10dB steps. The probe sound S2 was played at a fixed SPL immediately after S1. No delay was imposed between the offset of S1 (after 5-ms fall time) and the onset of S2 (before 5-ms rise time). Reductions in S2 responses were used to infer the strength of S1-evoked, forward suppression. To evoke reliable probe responses, S2 was chosen from BF-tone, narrowband noise (centered at BF, bandwidth ≤ 0.5 octaves) and broadband noise played at a preferred SPL of a neuron. For broadband noise, identical Gaussian noise tokens were used for S1 and S2. S1 and S2 each had a duration of 100 ms.
In experiments S1 and S2 were delivered from variable locations. S2 was always delivered from a loudspeaker in channel B (gray) at the best speaker location of a neuron, or from an adjacent location if the best speaker was not assigned to channel B. The S2 location was denoted as the “center” location. S1 was always delivered from a loudspeaker in channel A (black), which included speakers immediately adjacent to the “center” (denoted as the “near” location, <70° to the “center” location) and speakers far away from the “center” (denoted as the “far” location, >70° to the “center” location, often in the opposite hemifield), where the spatial separation indicates the angular distance between two speaker locations on a great circle. Additionally, S1 and S2 were summed physically (TDT, SM3) and played from the “center” location.

In experiments, the exact positions of “center”, “near”, and “far” varied from neuron to neuron, depending on the size and orientation of a SRF. As shown in Fig. 1B, the positions of speakers in channels A and B were interleaved. This arrangement was made to ensure that S2 in channel B (gray) could be played from either contralateral, midline, or ipsilateral locations. Moreover, for each S2 location used, S1 in channel A could be assigned to at least two “near” locations at the same hemifield to S2 and to at least three “far” locations at the opposite hemifield to S2. Since the two-stimulus stimulation had to be conducted after characterizing frequency and spatial selectivity of a neuron, not every “near” and “far” location were tested due to time constraints of single-unit recordings. We reported the results of neurons that have been tested at least one “near” and one “far” location along with the “center” location. When more than one “near” and “far” location was surveyed, the averaged results were reported for each spatial configuration. For control analyses, the S2-alone responses were measured in
each of the three spatial configurations and the S2-alone trials were randomly interwoven into (S1, S2) trials. S1-induced forward suppression was assessed relative to the S2-alone rate obtained in the same spatial configuration. This minimized the effects of potential drift in the overall excitability of a neuron on the observed strength of suppression.

**Identification of A1 and the caudal field**

Single-unit responses were collected from the primary auditory cortex (A1) and the caudal field in four hemispheres of one male and two female adult marmoset monkeys. In the marmoset, A1 is situated largely ventral to the lateral sulcus on the superior temporal plane and exhibits a low-to-high topographical frequency gradient along the rostral-caudal axis (Aitkin et al. 1986). Similar to other non-human primate species, the caudal medial (CM) and caudal lateral (CL) fields in marmoset can be identified by a tonotopic reversal with an abrupt decrease of BFs at the high-frequency border of A1 (Aitkin et al. 1986; Kaas and Hackett 2000; Merzenich and Brugge 1973).

In this study, a total of 681 tone-sensitive neurons were collected to construct the tonotopic maps of the three subjects.

Due to sampling limitations inherent in chronic recording, topographical mapping was accomplished with variable degrees of detail across subjects. The tonotopic organization in A1 and emergence of low-BF neurons in the caudal field were clearly seen in the auditory cortex of subjects M16s and M79u (Figs. 2A1 and 2B1), whereas tonotopic gradients were less continuous in A1 of subject M43s in both hemispheres (Fig. 2C1). To verify the location of the caudal field, we analyzed additionally the frequency selectivity of local field potentials (LFPs, 0.1~300 Hz) at individual recording
sites (~1mm craniotomy). The recording sites that were assigned to the caudal field all showed low-frequency profiles in LFPs and the presence of low-BF neurons. Since the medial to lateral division was not investigated in this study, we did not separate neurons further into CM and CL sectors.

Measurement of the acoustic properties of signals in marmoset ear canal

After the completion of electrophysiological experiments, the head-related impulse responses were measured in two subjects (M16s and M79u) in the awake condition using $2^{14}$-point Golay code stimulation (Zhou et al. 1992). Subjects sat normally in the primate chair. The pressure waveforms were collected using hearing aid microphones (Knowles Electronics model FG-23629-P16) inserted into the ear canals at a depth of 5-10 mm. The signals were amplified (40 dB) using a custom-built amplifier and digitized (100 kHz sampling rate) using the TDT system. In data analysis, the impulse responses were truncated by a 512-point Hamming window to obtain the direct responses with the maximal amplitude centered around ~2.56 ms. The time taken for the sound to travel from the loudspeaker to the position of the monkey head is ~2.3 ms. The power spectrum of the direct responses was described as the head-related transfer function (HRTF).

Interaural level difference (ILD) were extracted from frequency-domain signals, and expressed as the dB power difference between the right- and left-ear HRTFs averaged over a 2-kHz band. In the analysis, the left ear was designated as the ipsilateral ear, and positive ILDs corresponded to contralateral source locations. For controls, the free-field signals were collected at the positions of the animal’s ears (no monkey, two microphones pointing towards AZ -90° and 90°, respectively). Monaural
spectrum and ILD information for the free-field signals was compared with those of ear-
canal signals in the Results.

**Statistical significance tests**

The trend analysis on a given data set was based on a linear regression $t$-test; the $R^2$ and $t$ statistics of the slope were reported. Wilcoxon *rank-sum* tests were used to evaluate the population medians and two-sample *Kolmogorov-Smirnov* tests were used to evaluate the overall distributions between two groups of data sets. An alpha level of 0.05 was used for all statistical tests.
**Results**

**Spatial selectivity of A1 and CM/CL neurons to the frontal-field locations**

The data were obtained from studies of 406 single neurons (208 in A1 and 198 in CM/CL; a total of three marmosets) that responded significantly to broadband noise stimuli at one or more speaker locations (t-test, \( p < 0.05 \)). The 25th, 50th, and 75th percentiles of SPLs tested were 30 dB, 40 dB, 60 dB in A1 and 40 dB, 50 dB, 60 dB in CM/CL.

Figure 2A1-C1 show the tonotopic mapping in the auditory cortex of three monkeys. The border between area A1 and CM/CL was determined by the frequency reversal at the high-frequency end of A1. (See Methods for details on area identification.) The scatter plots show the best speaker locations of A1 neurons (Fig. 2A2-C2) and CM/CL neurons (Fig. 2A3-C3). A wide range of spatial selectivity was observed in both cortical areas. As summarized in Table I, a great portion of A1 and CM/CL neurons responded maximally to speakers at contralateral locations (~50%, \( AZ < 0^\circ, \text{contra} \)) and at locations above the horizontal plane (~60%, \( EL 45^\circ \)). Many neurons also showed preferences to ipsilateral (~25%, \( AZ > 0^\circ, \text{ipsi} \)) or midline (~25%, \( AZ = 0^\circ \)) locations. Space representation is thus not strictly lateralized by hemisphere in the marmoset auditory cortex. Further analyses revealed that the overall distributions of spatial preferences were rather similar among low-, mid-, and high-frequency neurons (Table II). We found no obvious correlation between spatial selectivity of a neuron (in \( AZ \) or \( EL \)) and its BF (\( R^2 < 0.01, p > 0.33 \) in A1; \( R^2 < 0.05, p \geq 0.04 \) in CM/CL).

------- Figure 2 here -------
Analysis of acoustic reflection and directional cues in the experimental setup

Since this study used free-field sound stimulation, reflections off the apparatus might interfere with the spatial information carried by the direct sound and thus affect a neuron’s spatial selectivity. Of particular concern were the near-field acoustic reflections off the top of the plexiglass primate chair and the stainless steel electrode manipulator (Fig. 1A). To address this concern, we tested free-field and ear-canal signals in response to golay-code stimulation with and without these two pieces of equipment after the completion of physiological experiments. Two subjects were used.

Figure 3 shows the results associated with two source locations (marked as black circles in Fig. 3A). Figure 3B compares ear-canal signals measured at the left ear (“ipsi”) of one subject (M79u) with and without the electrode manipulator. For the ipsilateral source location (left column), reflected waves were noticed in the received signal plotted in the time domain (arrow “a”, black curve). This reflection enlarged the notch depth in the corresponding HRTF. Among the fifteen locations tested, the results from this ipsilateral source location (EL 45°; AZ 60°) exhibited the greatest influence of reflection. It can be seen that reflected waves were much weaker in the received signal emitted from the contralateral location and the overall shape of the HRTF remained
relatively unchanged (right column). At both source locations the changes in HRTFs were more prominent at frequencies above 12 kHz.

To quantify these results, we measured the standard deviation (SD) of the difference between HRTF spectra with and without the manipulator for a given source location. The analyses were separately conducted in three frequency bands (2-12 kHz, 12-24 kHz, and 24-32 kHz) because they contained, respectively, the resonant peak, first notch (FN), and high-frequency features of marmoset HRTFs (Slee and Young, 2010). The results collected by the left-ear microphone are shown in Fig. 3C. Reflections caused spectral deviations of 1-6 dB. The most pronounced changes occurred over the FN range at EL45° (green dashed line). The results for the right-ear signals (not shown) exhibited similar patterns with smaller SD magnitudes.

Since the top plate of the primate chair was used to restrain the marmoset, it was impractical to remove the top plate and measure its effect on ear-canal signals. We therefore analyzed free-field signals, collected at the position of a monkey’s ear, with and without the top plate. The results collected by the left-ear microphone are shown in Fig. 3D. Similar to the results in Fig. 3B, reflected waves were more prominent in the signal received from the ipsilateral speaker at EL45° than from the contralateral speaker at EL0°. These reflections (left column, arrow “b”, black curve) caused accentuated spectral notches in the signal spectrum at 5-6 kHz, 25 kHz, and 32 kHz. Figure 3E plots the SDs of the difference between HRTF spectra (with and without the top plate) collected for a given source location. For results analyzed in the three frequency bands, SDs were larger for sources located at EL45° than at EL0°. Because sound waves emitted from EL45° had a larger angle of incidence than from EL0°, there was a greater
amount of sound energy reflected off the top plate. The results collected by the right-ear microphone (not shown) exhibited similar patterns and SD magnitudes.

These observations prompted us to further investigate the effects of the apparatus on the directionality of signals received at the ear canals. We reasoned that directly comparing the free-field (without monkey) and ear-canal (with monkey) signals would reveal the contributions of room acoustics to the directional filtering of HRTFs. Since the majority of neurons we collected had BFs greater than 2 kHz (Fig. 2), our analyses focused on two level-related localization cues: monaural spectrum and binaural ILDs. Results for free-field and ear-canal signals were compared in the frequency range from 2 to 32 kHz (the bandwidth of noise stimuli used in our experiments). The top plate was present in both conditions and results for two subjects are shown.

Figure 4A plots the monaural spectra of signals collected by the left-ear microphone for source locations at EL45° (top) and EL0° (bottom). The color of each grid represents the signal power at a given frequency averaged across a 2-kHz range. Two observations were made. (1) Free-field signals exhibited a spectral notch between 5-9 kHz across the AZ angles tested at EL45° (arrow a, top left; also see the example shown in Fig. 3D). This notch region was not consistently observed in the collected ear-canal signals (middle and right column). This discrepancy causes difficulties in identifying the sources of notches observed in ear-canal signals (see the example
shown in Fig. 3B) without the knowledge of the reflective and diffractive nature of body
and head of the monkey and that of the experimental setup. (2) The energies of free-
field signals at a given frequency showed no clear AZ dependence at either elevation. In
comparison, those of ear-canal signals were much enhanced at the resonant frequency
range (<12 kHz) and became increasingly directional toward midline-ipsilateral
directions between AZ0° and AZ60° at high frequencies (>12 kHz, arrow b). The
frequency dependent changes in the directionality of the cochlear microphonic or
monaural gain (re. free-field responses) has been described as the acoustic axis of
pinna (Middlebrooks and Pettigrew 1981; Phillips et al. 1982). For the two marmosets
tested, the maximal monaural energies became more sharply defined towards midline
at higher frequencies. Similar observations have been found in other animals (e.g., barn

The ILDs measured in these two conditions also showed clear differences (Fig.
4B). The ear-canal ILDs decreased in an orderly fashion from contralateral to ipsilateral
positions at all frequency bands. The dynamic range of ear-canal ILDs was larger for
high frequencies (e.g., ±20 dB at 17-19 kHz) than for low frequencies (e.g, ±10 dB at 3-
5 kHz). This is consistent with the previous study of marmoset HRTFs (Fig. 7 in Slee
and Young 2010). In comparison, the free-field ILDs at EL0° showed a weak ~5dB gain
(<9 kHz) towards the contralateral field due to the directionality of the microphones. At
high frequencies, ILDs were weak, showing no systematic AZ dependence at either
elevation.

The above analyses indicate that the AZ directionality of monaural spectrum and
ILD patterns observed in ear-canal signals was associated with HRTF filtering, not the
apparatus. However, reflections may influence the elevation selectivity of a neuron by altering the spectral profiles of HRTFs (Fig. 3). In this study, the effects of the acoustic features of the sound field on cortical SRFs were difficult to characterize because recording and restraining devices could not be removed during experiments (e.g., top plate of the primate chair). With this limitation in mind, the following experiments emphasized the relative changes in spatial tuning properties of a neuron as a function of sound level.

--- Figure 4 here ---

Effect of sound level on spatial tuning of individual neurons in A1 and CM/CL

Figure 5 shows the spatial responses of four example neurons measured at two SPLs (Fig. 5AB, A1 neurons; Fig. 5CD, CM/CL neurons). The raster plot, rate-azimuth function, and SRF of a neuron were analyzed at each SPL. The best area of a SRF is outlined in each SRF. Spatial responses of these neurons exhibited rich temporal patterns, including onset/offset (Fig. 5A) and sustained activity (Fig. 5C), consistent with the results reported in the awake cats (Mickey and Middlebrooks 2003) and awake macaque monkeys (Woods et al. 2006). Comparing results obtained at two different SPLs, the example SRFs could either expand (Figs. 5B, D) or contract (Fig. 5C) with increasing SPL. These level-induced modulations were not uniformly present in space and time. For example, response enhancement only occurred at the contralateral locations in Fig. 5B, and suppression was more evident during the onset phase of responses in Fig. 5A. The variability found in these individual SRFs contrasts with wide-
spread enhancement of onset activity found in anesthetized animals (e.g., Brugge et al. 1996).

--- Figure 5 here ---

To quantify the observed changes in cortical SRFs with increasing SPL, we analyzed the percent distributions of changes in best area ($\Delta$BA, Fig. 6AB), modulation depth ($\Delta$MD, Fig. 6CD), average firing rate across fifteen speaker locations ($\Delta$R_{avg}, Fig. 6EF), and the AZ and EL angle of best speaker location ($\Delta$AZ and $\Delta$EL, Fig. 6GH). Among these metrics, BA is a local measure of tuning based on responses around the preferred location, whereas MD and $R_{avg}$ are two global measures of tuning based on responses at both preferred and non-preferred locations. The measurement compared the response properties of a neuron between the lowest (SPL_{low}) and highest (SPL_{high}) sound levels tested, e.g., $\Delta$BA=$BA_{high}$-$BA_{low}$. A majority of neurons (>88%) were tested with SPL increment of 20dB or more. The 25th, 50th, and 75th percentiles of SPL_{low} were 30dB, 30dB, 40dB in A1 and 30dB, 40dB, 50dB in CM/CL. The percentiles of SPL increment (SPL_{high}-SPL_{low}) were 20dB, 20dB, 30dB in both A1 and CM/CL.

A total of 99 A1 neurons and 108 CM/CL neurons were tested. In both cortical areas, about 50% of neurons showed no variation in tuning acuity ($|\Delta$BA|$$\leq$$0.1; $|\Delta$MD|$$\leq$$0.1) and average firing rate ($|\Delta$R_{avg}|$$\leq$$5spk/sec), whereas the remaining ones showed either increased or decreased spatial sensitivity and excitability with increasing SPL. Moreover, many A1 and CM/CL neurons retained their spatial preferences (~30% with $|\Delta$AZ|=0°; ~50% with $|\Delta$EL|=0°), whereas the others showed contra/ipsi and
up/down shifts. These results show that increasing SPL evoked opposite changes in
tuning width, overall excitability, and tuning preference of individual neurons in the
auditory cortex of awake marmosets. The bidirectional modulation results in near-zero
medians for all metrics examined as shown in Fig. 6 (p>0.17 in A1; p>0.31 in CM/CL;
ranksum test).

Bidirectional changes in a tuning metric may yield a fixed group mean, but they
do not warrant a stable distribution of a tuning metric - divergence and convergence
could both occur. We next evaluated the level tolerance of spatial sensitivity of neurons
as an ensemble. The data set includes neurons showing in Fig. 6 along with neurons
tested at only one SPL. Figure 7 reports the overall distributions of best area,
modulation depth, and average firing rate at four SPL ranges. In A1 (filled circles), the
medians of BAs (Fig. 7A) were similar at low, moderate, and high SPLs (p>0.05;
ranksum test). As shown by the 25-75 percentiles of a data set (vertical bars), the
overall distributions of BAs were also preserved between low and high SPLs (p>0.14,
two-sample Kolmogorov-Smirnov test). Notably, narrowly and broadly tuned SRFs were
found at each SPL group. Similar observations were made in MD (Fig. 7B) and R_{avg}
(Fig. 7C). The medians and overall distributions of each metric were statistically
indistinguishable between low and high SPLs (p>0.05, ranksum test; p>0.25, two-
sample Kolmogorov-Smirnov test). At the lowest SPLs, responses showed slightly
narrower BAs (p<0.05) and higher R_{avg} (p<0.001) relative to those measured at higher
SPLs; *ranksum* test. In CM/CL (open circles), neurons had significantly smaller BAs and larger MD values than A1 neurons at moderate and high SPLs (*p*<0.05, *ranksum* test), but not at low SPLs. No between-group differences were detected for any of the three metrics in terms of median (*p*>0.09, *ranksum* test) and overall distribution (*p*>0.05, two-sample *Kolmogorov-Smirnov* test) across SPLs.

These results indicate that the spatial acuity and excitability of neurons as an ensemble were preserved over a large dynamic range of SPL in A1 and the caudal field of the awake marmosets. Further analyses revealed that reductions in *R*$_{avg}$ at higher SPLs were strongly correlated with increases in MD ($R^2$=0.37 in A1; $R^2$=0.25 in CM/CL; *p*<10$^{-10}$) and weakly correlated with decreases in BA ($R^2$=0.05 in A1; $R^2$=0.04 in CM/CL; *p*<0.05). This indicates that increasing sound intensity could suppress the responses of some cortical neurons, while improving their spatial tuning acuity.

----- Figure 7 here -----

While their spatial acuity was level tolerant, A1 and CM/CL neurons showed a wide-range of spatial selectivity to frontal-field sound locations (Fig. 2). Could the ensemble activities of these neurons also retain their spatial representation across SPL? In the previous study of the auditory cortex of anesthetized cat, the level-tolerant AZ coding was achieved through the disparity analysis between the responses of “contra” and “ipsi” channels (Stecker et al. 2005). Here we divided the neurons shown in Fig. 7 into three groups: “contra”, “midline”, and “ipsi” based on the AZ angles of their best speaker locations. For each neuron, we extracted its peak-normalized rate-AZ
tuning functions collected at EL0° and EL45°, and then calculated the ensemble average of the tuning functions of neurons within the same AZ and SPL group at these two elevations. Figure 8 shows the results of A1 (Fig. 8A) and CM/CL (Fig. 8B) neurons collected in four SPL ranges. The number of neurons included is given on each panel.

In A1, more “midline” neurons were observed relative to “contra” and “ipsi” neurons at very low SPL (≤20dB), whereas more “contra” neurons were observed at higher SPLs. In CM/CL, “contra” neurons dominated at all SPLs. Due to limited data points at the very low SPLs, we could not reliably assess the significance of this distinction. Comparing the results of three AZ groups (columns), the average tuning curves of “midline” neurons had a closed shape, whereas those of “contra” and “ipsi” neurons were half-open and peaked at lateral AZ angles. On average, the AZ tuning curves were modulated at a depth of 30-50% (the difference between the peak to trough of a curve) at both elevations and their overall shape did not change drastically between low and high SPLs (30-80dB). These results show that the ensemble responses of neurons stayed tuned across a large dynamic range of SPLs in the auditory cortex of marmoset monkeys. This applied to neurons with “contra”, “midline”, or “ipsi” preferences in both A1 and CM/CL.

Two-source stimulation revealed broadly distributed suppression in cortical SRFs

In comparison to results collected in the anesthetized condition (e.g., Brugge et al. 1996; Middlebrooks et al. 1998), one notable difference is that many neurons

-------- Figure 8 here --------
collected in the awake condition decreased their tuning widths and average firing rates with increasing SPL (Fig. 6). Intracellular studies have shown that the non-monotonic rate-level characteristics of cortical neurons are mediated by increasing strengths of synaptic inhibition at high sound levels (Tan et al. 2007). However, the spatial-location selectivity of synaptic inhibition is largely unknown in the literature.

Here we examined the level tuning of responses suppression at the excitatory center and far region of a SRF using a pair of lead-lag sounds (S1 and S2). The strength of suppression was inferred from reductions in neural responses to the lagging S2, i.e., forward suppression. Since S1 and S2 did not overlap in time, we reasoned that forward suppression would be solely attributed to properties of the leading S1, not to acoustic smearing or interaction between S1 and S2 at the ear. In experiments, the test stimulus S1 was 100-ms broadband noise, similar to those used for characterizing SRFs. The probe stimulus S2 was either 100-ms BF tone or BF-centered band-pass noise or broadband noise, whichever evoked significant excitatory responses. S1 was played from center and surround locations of a SRF (denoted as "center", "near", and "far" locations) from -10 to 80dB SPL and S2 was played immediately after the offset of S1 from the "center" location at a fixed SPL. (See details of the experimental design in Methods.)

Figure 9 shows the results of two example neurons. The first example was collected from CM/CL (the same neuron shown in Fig. 5D) whose SRF expanded with increasing sound level, and exhibited a broad spatial selectivity to frontal-field locations at 50dB (Fig. 9A). Figure 9B shows the raster plots of its response to sequentially presented S1 and S2. S1-evoked excitation (0-100ms, light gray) and forward
suppression (100-200 ms, dark gray, relative to the S2-alone response shown in the top row) can be seen at all three S1 locations tested (“center”, “near”, and “far”, marked in Fig. 9A). To examine the level dependence of S1-evoked excitation, we calculated the increase of the discharge rate during S1 relative to the spontaneous rate of the neuron (i.e., R(S1)-spont). For S1-evoked forward suppression, we calculated the reduction of discharge rates during S2 with and without the preceding S1 (i.e., R(S1,S2)-R(S2)). Figure 9C shows that the magnitudes of excitation and forward suppression generally increased with the S1 level. Note that excitatory responses were not always followed by suppression (e.g., responses to S1 at 30dB at “center” and “near” locations, Fig. 9C), suggesting that habituation of the postsynaptic response may not contribute to the observed forward suppression.

Forward suppression was also found in the absence of excitation. This is the case for the second example neuron, which responded exclusively to ipsilateral source locations (Fig. 9D). For the sequential presentation, the spiking activity to S1 peaked at 30dB SPL and then decreased to zero at higher S1 levels at both “center” and “near” locations (first and second columns, Fig. 9E), showing non-monotonic rate-level dependences. Interestingly, forward suppression persisted even when S1 did not elicit spiking activity between 60 and 80dB SPL. S1-evoked forward suppression is also seen at the “far” location showing very weak spiking activity at SPLs tested (third column, Fig. 9E). Because inhibitory synaptic events are almost exclusively triggered by stimulus onset (Scholl et al. 2010), the non-responsiveness of this neuron at high SPLs may be ascribed to sustained “silent” suppression evoked by S1, not to a lack of excitation or to separate offset-sensitive inhibitory input. Figure 9F plots the quantitative measurement
on these responses. Among neurons tested, 33% in A1 (23/70) and 46% in CM/CL (25/54) exhibited persistent “silent” suppression at “far” locations (re. R(S2), $p<0.05$, t-test) at one or more SPLs, despite no significant excitatory responses to S1 (re. spont. rate, $p>0.05$, t-test) between -10 and 80 dB SPL.

The results of these two example neurons show that the strength of forward suppression was more prominent at high S1 levels at both center and far regions of SRFs, despite different spatial extents of excitation in their SRFs and different characteristics of rate-level tuning of excitation. To evaluate the generality of these findings, we compared the general trends of level dependences of excitation and suppression across the three spatial configurations based on response of a population of neurons. Figure 10 shows the population average of the strengths of excitation and forward suppression measured in the three spatial configurations (n=70 in A1 and n=54 in CM/CL). The data show that the magnitudes of excitation (Figs. 10AC) at the “center” and “near” locations were much higher than those measured at the “far” locations at SPLs above 20dB in both A1 and CM/CL. Note that since the results were averaged across neurons, the shape of the rate-level tuning curve of individual neurons influences that of their average. In this study, a majority of neurons showed non-monotonic rate-level tuning to S1 played from the “center” location (70%, 49/70 in A1 and 72.2%, 39/54 in CM/CL). For these neurons, their rate-level tunings peaked at SPLs less than the highest SPL tested (80dB). In both cortical areas, the distributions of the so-called best
level peaked at 30-50dB SPL (31% in A1; 44% in CM/CL) and 80dB SPL (30% in A1; 28% in CM/CL). The non-uniform distribution of best level explained to some extents the two-peak profiles seen in the averaged tuning of excitation shown in Fig. 10AC. In contrast to excitation, the tuning of forward suppression increased with S1 level for the three spatial configurations in both A1 and CM/CL (Figs. 10BD). The monotonic increment of suppression in population average is consistent with the observations made in individual neurons (Figs. 9CF).

The contrast in the spatial extent of excitation and suppression also applied to their overall strengths measured based on the absolute firing rates of neurons. In the data analyses, we estimated the overall strength of S1-evoked excitation at a location by summing the firing rates in a tuning curve of a neuron, \( R(S1) \)-spont, between -10 and 80 dB SPL. To ensure that only valid response suppression was counted, we estimated the overall strength of forward suppression at a location by summing the rates in a tuning curve, \( R(S1,S2)-R(S2) \), which were significantly lower than zero (\( t \)-test, \( p < 0.05 \)). Results of all neurons collected at the same spatial configuration (e.g., “center”) were then averaged. This analysis allowed us to compare the overall levels of excitation and forward suppression across the spatial configurations.

Figure 10EF shows the population average of results at each of the three spatial configurations (mean±SEM). It can be seen that the strengths of excitation were much reduced at the “far” locations relative to those at the “same” and “near” locations (*\( p < 0.001 \) in A1, *\( p < 10^{-7} \) in CM/CL), whereas the strengths of forward suppression exhibited no significant differences among three configurations (\( p > 0.3 \) in A1 and \( p > 0.5 \) in CM/CL); ranksum test. Taken together, these data indicate that the cortical spatial...
responses are modulated by suppression mechanisms. While the strengths of excitation
are much reduced at the far regions of SRFs, the strengths of suppression increase with
stimulus level at both center and far regions of SRFs.

-------- Figure 10 here --------
Discussion

This study investigated spatial response properties of auditory cortex neurons in the awake marmoset. The results were evaluated in comparative terms across experimental conditions with respect to sound level and sound location. The main findings are three-fold. (1) Space representation is not strictly lateralized by hemisphere in the marmoset auditory cortex. Neurons in A1 and CM/CL fields showed a broad spatial selectivity to frontal field sound locations (Fig. 2). (2) The spatial tuning of individual neurons could either expand, or contract, or change little with increasing sound level (Fig. 6), whereas the spatial acuity and spatial tuning of neurons as an ensemble remained level tolerant (Figs. 7 and 8). (3) Although the strength of excitation was much reduced at the far regions of SRFs, the strength of suppression increased with sound level at both center and far regions of SRFs (Fig. 10). Together, these findings suggest that the spatial selectivity of neurons in the auditory cortex of marmosets is modulated by suppression mechanisms, which may play an important role in limiting the sizes of SRFs at high sound levels.

Spatial selectivity of neurons in marmoset auditory cortex and relation to previous work

Interaural timing and level differences and monaural spectral cues for sound localization are first extracted by neurons in the brainstem (reviewed by Irvine 1992; Oertel and Young 2004; Yin 2002; Young and Davis 2002). In the auditory cortex, change the spatial location of sound greatly modulate neural responses in multiple cortical areas (Middlebrooks et al. 2002). In this study, we used broadband noise stimuli (2 to 32 kHz) delivered in free field to study spatial sensitivity of cortical neurons.
According to marmoset head-related transfer functions (Slee and Young 2010), the main spatial cues within this frequency range are ILDs for encoding AZ and the spectral shape of the HRTF magnitude for encoding EL. As shown in Fig. 4B, the dynamic ranges of ILDs measured in this study were similar to those previously reported in marmosets (cf., Fig. 5, Aitkin and Park 1993; Fig. 7, Slee and Young 2010). In comparison with others mammalian species, the percentage of contralateral preferring neurons found in the awake marmoset (~50%, Table I) closely matched those reported in A1 of anesthetized cat (Rajan et al., 1990; Samson et al. 2000), but was lower than those reported in awake cat (~83%, Mickey and Middlebrooks 2003) and awake macaque (Woods et al. 2006). It is possible that contralateral neurons that peaked outside the frontal field were not properly counted in this study, and their numbers could be substantial as shown by the two previous awake studies, both of which sampled a full 360° azimuth plane (Mickey and Middlebrooks 2003; Woods et al. 2006). We are uncertain whether the directionality of pinna also contributes to the observed difference. Similar to those observed in marmoset (Fig. 4), the acoustic axis of the pinna points to the frontal ipsilateral sound field in both cat (Middlebrooks and Pettigrew 1981; Phillips et al. 1982) and macaque (Spezio et al. 2000).

Upward dominance of EL selectivity has been reported in area AES and A2 of anesthetized cats at near-threshold levels (Fig. 4, Xu et al. 1998), but relatively few studies have characterized elevation selectivity of neurons in the auditory cortex of awake animals. In this study more than 50% of A1 and CM/CL neurons showed preferences above the horizontal plane in response to broadband stimuli delivered from the upper frontal hemifield. Because reflections from the primate chair and recording
instrument we used influenced the amplitude and frequency of spectral notches of ear-
canal and free-field signals to varying degrees (Fig. 3), we are uncertain to what extent
the observed EL selectivity was caused by reflections in the apparatus. Previous HRTF
measurements in marmosets show that FN frequency monotonically increases with EL
at lateral source positions with AZ larger than 60°-80°. In the frontal field and on the
contralateral side, FN frequency shows constant or disorder relationships with AZ and
EL (cf. Fig. 9 in Slee and Young, 2010). This imposed further difficulties in interpreting
the observed EL selectivity in the frontal field based on FN frequency, especially when
speculated spectral artifacts caused by acoustic reflections of the apparatus were not
consistently observed in ear-canal signals (Figs. 4A).

Mechanisms of level invariant space coding in auditory cortex

Although lesion studies have shown that spatially oriented behaviors during
sound localization require an intact auditory cortex (Heffner and Heffner 1990; Jenkins
and Merzenich 1984; Thompson and Cortez 1983), the reliability of spatial functions of
the auditory cortex is not fully understood. One major issue is that focal representations
of space by cortical neurons are mostly observed at near-threshold levels in
anesthetized preparations (Brugge et al. 1996; Middlebrooks et al. 1998), reminiscent of
those observed in the inferior colliculus (Semple et al. 1983). At high stimulus levels,
SRFs generally expand in width and their spatial acuity and boundaries can no longer
be reliably defined, in contrast to level-robust perceptual performance (Sabin et al.
2005). Stecker and colleagues propose that this level issue could be resolved by using
disparity information created between two “opponent” neural populations composed of
contralateral and ipsilateral units within each hemisphere. Their analyses show that
although the tuning of individual populations broadens with increases of stimulus level, the difference in tuning between two populations remains unchanged, providing sufficient information about source azimuth at both low and high SPLs (Stecker et al. 2005).

Spatial responses measured in the awake condition differed markedly from those measured in the anesthetized condition. We found that spatial tuning widths of individual neurons in A1 and CM/CL of awake marmosets do not obey a fixed relationship with stimulus level - expansion and contraction of SRFs were both observed along with those showing no change (Fig. 6). These results are consistent with findings in the auditory cortex of awake cats (Mickey and Middlebrooks 2003). Additionally, the distribution of spatial tuning width remained unchanged between low and high SPLs (Fig. 7A), consistent with findings in the auditory cortex of awake macaques (cf. Fig. 8, with the exception of one subject, in Woods et al., 2006). Similarly, the average firing rates of neurons as a population changed little between low and high SPLs (Fig. 7C), whereas that of individual neurons could either increase or decrease (Fig. 6EF).

These results depict an interesting relationship between the variability of responses of individual neurons versus the stability of their ensemble responses in the auditory cortex. In this regime, the changes in the spatial sensitivity of individual neurons are not random - SRF expansion is offset by SRF contraction with a zero net gain (i.e., near-zero medians in Fig. 6). As such, the ratio of narrowly and broadly tuned SRFs remains roughly the same across SPLs (as shown in percentiles in Fig. 7A). In our analysis, the averaged AZ tuning curves of contralateral, midline, and ipsilateral
preferring neurons remained tuned between 30-80 dB SPLs (Fig. 8), indicating that level-tolerant AZ responses are not limited to hemispherical channels.

One apparent limitation of this study is that the sampling space is not complete. The sampling issue has been recently addressed by Kuwada and colleagues in studying the azimuth coding in the inferior colliculus of unanesthetized rabbit (Kuwada et al. 2011). They report that the top 35% of neurons showed level-tolerant tuning within ± 150° ranges of azimuths. It remains to be tested whether the population activity in the auditory cortex shows level-tolerant tuning in front/back and up/down dimensions and whether the general principles of the opponent-channel theory promote level invariant space coding along specific spatial dimensions.

**Suppression mechanisms underlying cortical spatial selectivity**

In contrast to previous findings in the anesthetized condition (e.g., Brugge et al. 1996; Middlebrooks et al. 1998), neurons collected in the awake condition could decrease their tuning widths and excitability with increasing SPL (Fig. 6 in this study and Fig. 13 in Mickey and Middlebrooks 2003), suggesting the involvement of inhibitory/suppressive activity in cortical SRFs. This study investigated the suppression mechanisms using a modified “forward-masking” paradigm (Brosch and Schreiner 1997; Calford and Semple 1995). We did not estimate the strength of suppression during the presentation of the test stimulus (by simultaneously presenting S1 and S2) due to concerns about acoustic interactions of sound waves at the ear canal, as they could evoke the perception of “summing localization” (Blauert 1997). Results show that although excitatory spiking activity is substantially reduced at the far regions of SPFs, the level dependence and the overall strength of forward suppression were similar in
results obtained at center and far regions of SRFs (Fig. 10). This observation suggests that suppressive activity on average is more broadly tuned than excitatory activity in cortical SRFs. Another noteworthy feature of our results is the “silent” suppression observed at some SPLs and/or locations (Fig. 9D). Because inhibitory synaptic events are almost exclusively triggered by stimulus onset (Scholl et al. 2010), it is unlikely that the observed forward suppression was driven by separate offset-sensitive inhibitory input. One plausible explanation is that the observed “silent” suppression was governed by inhibitory synaptic activity, which reduced the sizes of SRFs at certain sound levels.

Nevertheless, the strength of suppression measured in extracellular studies is not equivalent to that of synaptic inhibition and could be affected by membrane adaptation mechanisms. A complete survey of the structures of cortical SRFs requires knowledge of the spatial tuning of synaptic excitation and inhibition using intracellular techniques. Although intracellular work has revealed prominent inhibitory activity underlying frequency (Kaur et al. 2004; Tan et al. 2004; Wehr and Zador 2003; Zhang et al. 2003) and intensity (Tan et al. 2007; Wehr and Zador 2003) selectivity of neurons in A1 of anesthetized rat, information about spatial tuning profiles of synaptic inputs is limited in the literature (Chadderton et al. 2009) and synaptic inhibitory activity in cortical SRFs has not been reported. It is unknown whether the feature selectivity of cortical inhibition differs between spatial and non-spatial aspects of sound analyses.

**Comparison of spatial processing in A1 and the caudal field**

Based on the differential distributions of spatial and spectral tuning widths and monkey call selectivity among cortical areas (Rauschecker et al. 1995; Recanzone 2000; Tian et al. 2001), researchers suggest that the central auditory system is divided
into at least two separate streams, the rostral-ventral ("what") and caudal-dorsal
("where") pathways emerging after A1 (Kaas and Hackett 2000; Recanzone and Cohen
2010; Romanski et al. 1999). It has been proposed that spatial information of sound is
further analyzed and better represented by neurons in the caudal field in primate
species.

In this study, the caudal-field neurons exhibited better spatial acuity than A1
neurons (Fig. 7) in agreement with previous findings in the awake macaques (Woods et
al. 2006). Although these results favor the hypothesis that the caudal field is more
suited to space coding, spatial sensitivity of CM/CL neurons do not differ substantially
from that of A1 neurons. In particularly, the overall distributions of BA, MD, and $R_{avg}$
were markedly similar between the two cortical fields (Fig. 7) and strong suppressive
modulations are found in SRFs of both A1 and CM/CL neurons (Fig. 10).

Whether spatial sensitivity of neurons in A1 and the caudal field would differ in an
active listening situation is not addressed in this study. Compelling anatomical and
physiological evidence indicates that the caudal field is a site of multisensory
convergence, where visual influences are mostly suppressive at the level of single
neurons (de la Mothe et al. 2006; Kayser et al. 2009; Schroeder and Foxe 2002; Smiley
et al. 2007). It remains to be tested whether suppression as shown in this study serves
to improve space coding in a spatial task involving multisensory experience.
Acknowledgements:

This work was supported by NIH Grants DC 03180 and DC 005808. We thank J. Estes and N. Sotuyo for assistance with animal care and E. Issa, S. Slee, E. Young, M. Jeschke, D. Gamble and E. Remington for discussions on results. B. May and S. Slee have provided invaluable expertise on HRTF measurement in marmoset. Special gratitude to N. Sotuyo for providing drawings of the experimental setup.
**Reference:**


Thompson GC, and Cortez AM. The inability of squirrel monkeys to localize sound after unilateral ablation of auditory cortex. *Behav Brain Res* 8: 211-216, 1983.


### Tables

<table>
<thead>
<tr>
<th></th>
<th>A1 (n=208)</th>
<th>CM/CL (n=198)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0°</td>
<td>57</td>
<td>63</td>
</tr>
<tr>
<td>45°</td>
<td>122</td>
<td>121</td>
</tr>
<tr>
<td>90°</td>
<td>29</td>
<td>14</td>
</tr>
<tr>
<td>Azimuth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>contra</td>
<td>95</td>
<td>103</td>
</tr>
<tr>
<td>midline</td>
<td>58</td>
<td>45</td>
</tr>
<tr>
<td>ipsi</td>
<td>55</td>
<td>50</td>
</tr>
</tbody>
</table>

**Table 1: Elevation and azimuth selectivity of neurons in A1 and CM/CL.** For elevation tuning, the total numbers of neurons were counted based on the EL angles of their best speaker locations. For azimuth tuning, the contralateral (contra) group included neurons that responded maximally to speakers at AZ -30°, -60°, and -90°; the midline group included neurons that responded maximally to speakers at AZ 0° and the top speaker; the ipsilateral (ipsi) group included neurons that responded maximally to speakers at AZ 30°, 60°, and 90°.
Table 2: Spatial selectivity of neurons in three BF ranges. The grouping criteria for elevation and azimuth selectivity were identical to those shown in Table I. The analyses were restricted to neurons whose BFs were identified using pure-tone stimuli (156 in A1 and 127 in CM/CL). The 25th, 50th, and 75th percentiles of BFs were 4.92 kHz, 10.12 kHz, and 14.6 kHz, respectively, in A1 and 4.59 kHz, 14.93 kHz, and 23.36 kHz, respectively, in CM/CL.
Figure legend

**Figure 1: Experimental setup in the acoustic chamber.** (A) Schematic illustration of the rear view of the acoustic chamber. The subject sat in the primate chair and an electrode manipulator was placed beside the animal. Fifteen loudspeakers were positioned to cover the frontal hemifield of the room. The wall, loudspeaker frame, and recording table were covered with acoustically absorptive foam to reduce reflections. (B) The frontal view of the loudspeaker positions was plotted in spherical coordinates. (Contra, contralateral; ipsi, ipsilateral.)

**Figure 2: Tonotopy and best speaker distributions in area A1 and CM/CL of three subjects.** (A1-C1) Tonotopy obtained from four hemispheres in three subjects. Dashed lines mark the tentative division between A1 and CM/CL. Gray line marks the position of the lateral sulcus (LS), which was visually identified on the skull lateral to the coronal suture during implant surgery. Colored dots indicate the BF of individual neurons. When more than one neuron was encountered on an electrode track, their positions were scattered by ¼ mm for visual inspection. (A2-C2) Scatter plots of the best speaker locations of A1 neurons. (A3-C3) Scatter plots of the best speaker locations of CM/CL neurons. The radius of a circle is proportional to the number of neurons preferring a given speaker. When spatial responses were tested at multiple sound levels for a neuron, only the best speaker location of a SRF showing the smallest best area (BA) was reported. BA is defined in Methods.
**Figure 3: Analysis of acoustic reflections off the primate chair and electrode manipulator.** (A) Schematic illustration of the positions (EL, AZ) of two loudspeakers. (B) Effects of reflection off the manipulator on the impulse responses (top) and HRTFs (bottom) of ear-canal signals collected in the left ear. The stimuli were delivered for the two source locations shown in (A) and the manipulator was placed at a distance of ~13 cm to the left ear canal. (C) SDs of the difference between HRTF spectra with and without the manipulator. The data were analyzed in three frequency ranges and plotted as a function of the AZ angles of the fifteen loudspeakers. (D) Effects of reflection off the top plate of the primate chair on the impulse responses (top) and HRTFs (bottom) of free-field signals. The stimuli were delivered for the two source locations shown in (A). The top plate was placed at a depth of ~2.5 cm below the ear canal. The recording microphone was placed at the position of the left ear canal of the subject. (E) SDs of the difference between signal spectra collected with and without the top plate. The data were presented in the same format as those in (C).

**Figure 4: Analysis of the directional cues in free-field and ear-canal signals.** (A) Monaural spectra of free-field and ear-canal signals measured in the left ears of two subjects at EL0° and EL45°. The results report the magnitudes of the free-field spectra and HRTFs from 2 to 32 kHz. Each grid shows the average power over a 2-kHz band. (B) ILDs of free-field and ear-canal signals measured for the same subject. The ILDs were computed as the difference between the monaural spectra of right- and left-ear signals and each grid shows the average ILD over a 2-kHz band. All data were plotted...
as a function of the AZ of a loudspeaker location. The left ear was designated as the
ipsilateral ear, and positive ILDs corresponded to contralateral source locations.

**Figure 5: Effects of sound level on spatial selectivity of example cortical neurons.**
Each panel shows from top to bottom the raster plot, rate-azimuth tuning curve, and
SRF of a neuron in response to 100-ms Gaussian broadband noise played at two SPLs.
The data in the raster plots are organized based on the AZ angles of the loudspeakers;
asterisks mark those at EL 45°. The gray shading marks the duration of a stimulus and
the SPL used is indicated above each raster plot. The dashed line on a rate-azimuth
tuning curve indicates the spontaneous rate of a neuron. Best area (BA) was calculated
as the ratio between the SRF area above 62.4% peak rate (outlined in blue) and the
total area covered by the speaker array.

**Figure 6: Variation in spatial sensitivity of neurons with increasing sound level.**
Histograms show the distributions of changes in neuron best area (A)(B), modulation
depth (C)(D), average firing rate across fifteen locations (E)(F), and AZ and EL angles
of best speaker location (G)(H). Results of A1 neurons are shown on the left and those
of CM/CL neurons are shown on the right. Median values are indicated in each panel.
The SPL increment and SPL range used are given in the main text.

**Figure 7: Level dependence of spatial acuity and overall excitability of neurons in
A1 and CM/CL.** Best area (A), modulation depth (B), and average firing rate (C) are
reported as a function of SPL. In each panel, the gray dots show the original data points
and the vertical bars mark the 25th-75th percentile of a data distribution. Filled and open circles show the median values of results in A1 and CM/CL, respectively. The numbers of neurons in the four SPL groups (from lowest to high) were 31, 139, 108, and 78 in A1 and 18, 119, 139, and 75 in CM/CL.

**Figure 8: Ensemble average of AZ tuning in A1 and CM/CL.** (A)(B) Average AZ tuning curves of neuronal groups preferring contralateral, midline, and ipsilateral directions in A1 and CM/CL. The data set is identical to that shown in Fig. 7. Results from different SPL ranges are shown in different rows. Black, AZ tuning at EL0°; red, AZ tuning at EL45°.

**Figure 9: Response of two example neurons to sequential two-source stimulation.** (A) Raster plot and SRF of a CM/CL neuron in response to broadband noise (50dB SPL). (B) Raster plots of the responses of this neuron to two sequentially presented sounds, S1 and S2. S1 was a 100-ms broadband noise played from the “center”, “near”, or “far” location, as marked in (A). S2 was a 100-ms broadband noise played from the “center” location. The S1 level varied from -10 to 80dB SPL and the S2 level remained fixed at 50dB SPL. The duration of S1 and S2 are marked by light and dark gray, respectively. The top row in each raster plot shows the S2-alone response. (C) Rate responses during S1 (after subtracting the spontaneous rate), R(S1)-spont, were used to describe the level dependence of S1-evoked excitation. Rate responses after S1 (after subtracting the S2-alone rate), R(S1,S2)-R(S2), were used to describe the level dependence of S1-evoked suppression. (D)(E)(F) Spatial tuning and responses of an
A1 neuron to sequential stimulation. S2 was a broadband noise played at 30dB SPL. The data are presented in the same format as those in A-C.

**Figure 10: Level dependence of excitation and suppression at center and far regions of SRFs.** Population average shows the peak normalized rate-level functions of S1-evoked excitation (A)(C) and forward suppression (B)(D) for the three spatial configurations (mean±SEM). The results of A1 and CM/CL neurons are shown in the left and right columns, respectively. For S1-evoked excitation, the normalization factor was the maximal firing rate among the rate-level curves, R(S1)-spont, collected at “center”, “near”, and “far” locations. One normalization factor for excitation was used per neuron. For S1-evoked forward suppression, the normalization factor was the maximal absolute rate among the rate-level curves, R(S1,S2)-R(S2), collected at “center”, “near”, and “far” locations for each neuron. One normalization factor for suppression was used per neuron. (E)(F) Population average of the overall strengths of S1-evoked excitation and forward suppression for the three spatial configurations (mean±SEM).
Figure 1: Experimental setup in the acoustic chamber.
Figure 2: Tonotopy and best speaker distributions in area A1 and CM/CL of three subjects.
Figure 3: Analysis of acoustic reflections off the primate chair and electrode manipulator.
Figure 4: Analysis of the directional cues in free-field and ear-canal signals.
Figure 5: Effects of sound level on spatial selectivity of example cortical neurons.
Figure 6: Variation in spatial sensitivity of neurons with increasing sound level.
Figure 7: Level-invariant spatial acuity and overall excitability of neurons in A1 and CM/CL.
Figure 8: Ensemble average of AZ tuning in A1 and CM/CL.
Figure 9: Response of two example neurons to sequential two-source stimulation.
Figure 10: Level dependence of excitation and suppression at center and surround locations.