

Spectral Organization of the Human Lateral Superior Temporal Gyrus Revealed by Intracranial Recordings

Kirill V. Nourski¹, Mitchell Steinschneider^{3,4}, Hiroyuki Oya¹, Hiroto Kawasaki¹, Robert D. Jones² and Matthew A. Howard¹

¹Department of Neurosurgery, ²Department of Neurology, The University of Iowa, Iowa, IA 52242, USA ³Department of Neurology, ⁴Department of Neuroscience, Albert Einstein College of Medicine, New York, NY 10461, USA

K.V.N. and M.S. contributed equally to this work.

Address correspondence to: Kirill V. Nourski, Department of Neurosurgery, The University of Iowa, 200 Hawkins Dr. 1815 JCP, Iowa, IA 52242, USA. Email: kirill-nourski@uiowa.edu

The place of the posterolateral superior temporal (PLST) gyrus within the hierarchical organization of the human auditory cortex is unknown. Understanding how PLST processes spectral information is imperative for its functional characterization. Pure-tone stimuli were presented to subjects undergoing invasive monitoring for refractory epilepsy. Recordings were made using high-density subdural grid electrodes. Pure tones elicited robust high gamma event-related band power responses along a portion of PLST adjacent to the transverse temporal sulcus (TTS). Responses were frequency selective, though typically broadly tuned. In several subjects, mirror-image response patterns around a low-frequency center were observed, but typically, more complex and distributed patterns were seen. Frequency selectivity was greatest early in the response. Classification analysis using a sparse logistic regression algorithm yielded above-chance accuracy in all subjects. Classifier performance typically peaked at 100–150 ms after stimulus onset, was comparable for the left and right hemisphere cases, and was stable across stimulus intensities. Results demonstrate that representations of spectral information within PLST are temporally dynamic and contain sufficient information for accurate discrimination of tone frequencies. PLST adjacent to the TTS appears to be an early stage in the hierarchy of cortical auditory processing. Pure-tone response patterns may aid auditory field identification.

Keywords: classification analysis, electrocorticography, high gamma activity, human auditory cortex, pure tones

Introduction

The functional organization of the human auditory cortex is incompletely understood. This deficiency hampers our ability to clarify the mechanisms underlying the normal cortical processing of speech and other complex sounds, which is a prerequisite to thoroughly understand the pathophysiology of certain speech and hearing disorders. A fundamental organizational feature of classical auditory pathways is the orderly representation of sound frequency (Rose et al. 1963; Aitkin and Webster 1971). The auditory cortex in Old World monkeys is characterized by orderly frequency gradients that help define specific fields, both within the core regions on the superior temporal plane and along adjacent belt regions, including those on the lateral surface of the superior temporal gyrus (STG; Merzenich and Brugge 1973; Morel et al. 1993; Kaas et al. 1999; Rauschecker and Tian 2004; Tian and Rauschecker 2004; Romanski and Averbeck 2009).

While maps of frequency representation are used extensively in experimental animals to gain insights into the organization of auditory cortical fields, obtaining comparable data

in human subjects is quite challenging. Noninvasive methods such as functional magnetic resonance imaging (fMRI) and magnetoencephalography have limitations in their capacity to resolve cortical activity in the time and space dimensions, and direct recordings can only be obtained in neurosurgical patients who require the placement of electrodes as part of their clinical treatment plan. Such direct recordings from the posteromedial Heschl's gyrus (HG), the putative location of primary auditory cortex in the human (Hackett et al. 2001), have demonstrated spectral specificity of neuronal responses, but have not examined in detail the spatial extent of orderly frequency representations (Howard et al. 1996; Bitterman et al. 2008). fMRI studies have promoted several views concerning tonotopicity, with mirror-image representations described as oriented along the long axis of the HG (Formisano et al. 2003) or more perpendicular to the HG (Talavage et al. 2004; Woods et al. 2009, 2010; Humphries et al. 2010; Da Costa et al. 2011; Striem-Amit et al. 2011; Langers and van Dijk 2012).

The representation of sound frequency in the non-primary auditory cortex located on posterolateral superior temporal (PLST) gyrus has been even more difficult to define, leading to contradictory interpretations. Results from fMRI studies have ranged from an absence of spatially organized tonotopic gradients (Formisano et al. 2003; Lewis et al. 2009; Woods et al. 2009, 2010), to small gradients on the portions of PLST (Talavage et al. 2004; Humphries et al. 2010), to large-scale, mirror-image tonotopic patterns (Striem-Amit et al. 2011). Clarifying this issue is especially important given that this region of auditory cortex is crucial for acoustic-to-phonetic transformation of speech (Hickok and Poeppel 2004) and auditory scene analysis (Goll et al. 2010; Schadwink and Gutschalk 2010, 2011; Shamma and Micheyl 2010; Kashino and Kondo 2012). In the monkey, tonotopicity-based field delineation has contributed to a framework, wherein distinct areas are postulated to preferentially represent sound attributes such as location, identity, and relevance for communication (Recanzone et al. 2000; Tian et al. 2001; Rauschecker and Tian 2004).

Intracranial recordings demonstrate that PLST is responsive to a wide range of acoustic stimuli, including clicks, tones, and speech (Howard et al. 2000; Crone et al. 2001; Edwards et al. 2005; Chang et al. 2010; Steinschneider et al. 2011; Pasley et al. 2012) and is modulated by attention (Besle et al. 2011; Gomez-Ramirez et al. 2011; Mesgarani and Chang 2012) and visual input (Reale et al. 2007). Limited understanding of how PLST is organized hampers the development of conceptual frameworks for clarifying how these diverse

sounds are represented in the auditory cortex. To better understand how this acoustically responsive region of the cortex may relate to multifield models of the human auditory cortex, it is important to determine how basic spectral acoustic information is represented within PLST. The results of earlier human intracranial and fMRI studies demonstrate that PLST responds robustly to acoustically complex stimuli, including speech. Less is known about how this region of cortex responds to less spectrally complex stimuli, such as pure tones, and whether patterns of basic functional organization can be discerned based on responses to these stimuli (Obleser et al. 2007; Chang et al. 2010; Leaver and Rauschecker 2010; Steinschneider et al. 2011; Pasley et al. 2012).

In this report, we present our findings from experiments performed to study how pure-tone stimuli are represented within PLST, both across cortical space and time. Recordings were obtained from electrode arrays implanted in neurosurgical patients. We demonstrate that PLST cortical responses contain stimulus frequency-specific information, which changes over time and is largely unaffected by sound intensity.

Materials and Methods

Subjects

Experimental subjects were 13 neurosurgical patients (7 male, 6 female, age 20–56 years old, median age 33 years old). The subjects had been diagnosed with medically refractory epilepsy and were undergoing chronic invasive electrocorticography (ECoG) monitoring to identify potentially resectable seizure foci. Research protocols were approved by the University of Iowa Institutional Review Board and by the National Institutes of Health. Written informed consent was obtained from each subject. Participation in the research protocol did not interfere with acquisition of clinically required data. Subjects could rescind consent at any time without interrupting their clinical evaluation.

The patients initially remained on their antiepileptic medications, but were typically weaned from these drugs during the monitoring period at the direction of the patients' treating neurologists. Experimental sessions were suspended for at least 3 h if a seizure occurred. Further, the patient had to be alert and willing to participate for the research activities to resume.

Eleven subjects had left-hemispheric language dominance, as determined by intracarotid amygdala (Wada) test results; subject L162 had bilateral, and R139 had right language dominance. In 6 subjects, the electrodes were implanted on the left side, while in 7 others recordings were from the right hemisphere. The side of implantation is indicated by the letter prefix of the subject code (L for left and R for right). The hemisphere of recording was language dominant in 8 subjects (L130, L138, R139, L146, L151, L162, and L178) and nondominant in 6 subjects (R136, R142, R153, R175, R180, and R186).

All subjects underwent audiometric and neuropsychological evaluation before the study, and none were found to have hearing or cognitive deficits that could impact the findings presented in this study (median full-scale intelligence quotient = 97, range 77–115). All subjects were native English speakers. Intracranial recordings revealed that the auditory cortical areas on the STG were not epileptic foci in any of the subjects.

Experiments were carried out in a dedicated electrically shielded suite in The University of Iowa General Clinical Research Center. The room was quiet, with lights dimmed. Subjects were awake and reclining in a hospital bed or an armchair. Stimuli were presented in a passive-listening paradigm, without any task direction. Typically, a recording session was 10–20 min in duration. After each recording session, conversation with the subject ensued to confirm their awake state and willingness to continue with another experimental protocol.

Stimulus Presentation

Experimental stimuli were tones presented at 6 frequencies between 250 and 8 kHz in 1-octave steps (300 ms duration, 5-ms rise-fall time, 2-s interstimulus interval). An interleaved presentation paradigm was used, where the 6 tones were presented 50 times each in a random order. The stimuli were presented at a comfortable level, approximately 50 dB above hearing threshold. This paradigm was used in all subjects except R175. In subject L162, this experiment was repeated 4 times with the stimuli presented at 4 different intensities (26, 41, 56, and 71 dB SPL).

In 4 subjects (R175, L178, R180, and R186), an additional, extended stimulus set was also used to examine the degree in which response patterns were influenced by stimulus intensities. In this set, 12 tone frequencies between 250 and 8 kHz were presented at 6 different levels (11–61 dB SPL) in 10 dB steps in a random interleaved sequence. Tone duration, rise-fall time, and interstimulus interval in this set were 200 ms, 5 ms, and 1.2 s, respectively. Each stimulus was presented 24 times. In subject R175, only this extended stimulus set was presented. Because of the limited number of trials (24) at any given intensity, we pooled together trials corresponding to the 2 highest tone intensities (51 and 61 dB SPL) for each of the 6 frequencies used in the main experimental set for classification analysis (see below).

Stimuli were delivered via insert earphones (ER4B, Etymotic Research, Elk Grove Village, IL, United States of America) that were integrated into custom-fit earmolds. Stimulus delivery and data acquisition were controlled by a TDT RX5 or RZ2 real-time processor (Tucker-Davis Technologies, Alachua, FL, United States of America) with a sampling rate of 24 414 Hz.

Response Recording

Details of electrode implantation have been described previously (reviewed in Howard et al. 2012). In brief, ECoG data were recorded from multicontact subdural grid electrodes (AdTech, Racine, WI, United States of America) placed over the perisylvian cortex. The recording arrays consisted of platinum-iridium disc electrodes (2.3 mm diameter, 5-mm interelectrode distance), embedded in a silicon membrane. The electrodes were arranged in an 8 × 12 grid, yielding a 3.5 × 5.5 cm array of 96 contacts. A subgaleal contact was used as a reference. Electrode grids were placed solely on the basis of clinical requirements and were part of a more extensive set of recording arrays meant to identify seizure foci. Consequently, the number of auditory-responsive recording sites varied across subjects. Recording electrodes remained in place for a period ranging from 7 to 28 days (median 14 days) under the direction of the patients' treating neurologists. The ECoG data were amplified, filtered (0.7–800 Hz bandpass, 12 dB/octave rolloff), digitized at a sampling rate of 2034.5 Hz, and stored for subsequent offline analysis.

Subjects underwent whole-brain high-resolution T_1 -weighted structural MRI (resolution 0.78 × 0.78 mm, slice thickness 1.0 mm, 2 volumes for averaging) scans before electrode implantation to locate recording contacts. Averaging 2 volumes improved the signal-to-noise ratio of the MRI data sets and minimized the effects of movement artifact on image quality. Preimplantation MRIs and postimplantation thin-sliced volumetric computed tomography (CT) scans (resolution 0.51 × 0.51 mm, slice thickness 1.0 mm) were coregistered using a 3-dimensional rigid fusion algorithm (Analyze version 8.1 software, Mayo Clinic, MN, United States of America). Coordinates for each electrode contact obtained from postimplantation CT volumes were transferred to preimplantation MRI volumes. Results were compared with intraoperative photographs to ensure reconstruction accuracy.

Data Analysis

ECoG data obtained from each recording site were downsampled to a rate of 1000 Hz and analyzed in the time–frequency plane as event-related band power (ERBP). Prior to the calculation of ERBP, individual trials were screened for possible contamination from electrical interference, epileptiform spikes, high-amplitude slow-wave activity, or movement artifacts. To that end, individual trial waveforms

with voltage peaks or troughs greater than 2.5 standard deviations from the mean of the 50 stimulus presentations were rejected from further analysis. To minimize contamination with power-line noise, ECoG waveforms were notch-filtered at 60 Hz (10th order Chebyshev Type II filter, stopband 58–62 Hz, stopband ripple 2 dB down from the peak passband value). Data analysis was performed using custom software written in MATLAB Version 7.12.0 programming environment (MathWorks, Natick, MA, United States of America).

Time–frequency analysis of the ECoG was performed using wavelet transforms based on complex Morlet wavelets following the approach of Oya et al. (2002). Center frequencies ranged from 20 to 200 Hz in 5 Hz increments. ERBP was calculated for each frequency on a trial-by-trial basis and normalized to median baseline power, measured using wavelets centered between 200 and 100 ms prior to the stimulus onset. ERBP values were then log-transformed and averaged across trials.

Quantitative analysis focused on the high gamma ECoG frequency band, which has been shown to be a sensitive and specific indicator of auditory cortical activation (Crone et al. 2001; Brugge et al. 2009; Edwards et al. 2009; Steinschneider et al. 2008, 2011; Mesgarani and Chang 2012; Pasley et al. 2012). High gamma band was defined in the present study as the range of center frequencies between 75 and 150 Hz.

ERBP changes were first averaged within the following 3 bands: 75–90, 95–115, and 125–150 Hz, followed by averaging across the 3 bands. The purpose of binning together 5-Hz wide frequency bands into these larger bands was to avoid the over-contribution that the highest frequencies (125–150 Hz) would have if the 5-Hz wide bands were simply averaged together across the entire range of interest.

The wavelet constant ratio used for time–frequency analysis was defined as $f_0/\sigma_f=9$, where f_0 is the center frequency of the wavelet, and σ_f is its standard deviation in frequency. For $f_0=75$ Hz, this yielded a wavelet with a standard deviation in the frequency of 8.33 Hz and in the time of 19.1 ms. The wavelet with $f_0=150$ Hz was characterized by a standard deviation in frequency and time of 16.7 Hz and 9.55 ms, respectively. Contribution of energy from the post-stimulus onset interval to the estimate of baseline power—the “edge effect”—was negligible for the range of center frequencies that corresponded to the high gamma frequency band.

For display purposes, average high gamma ERBP changes in response to a given tone frequency within a given 50-ms wide time bin across all 8×12 grid contacts were smoothed using triangle-based cubic interpolation with an upsampling factor of 16 to produce cortical activation maps.

Average ERBP values were calculated for each recording site within 50-ms time windows with a 50% overlap, from 100 ms before stimulus onset to 725 ms after stimulus onset. Statistical significance of changes in responses from the baseline were determined via paired *t*-tests comparing average ERBP values within each time window versus that within an equal duration reference interval spanning 175–125 ms before stimulus onset. Correction for multiple comparisons was done by controlling the false discovery rate following the method of Benjamini and Hochberg (1995) and Benjamini et al. (2001) ($q < 0.01$).

Activation patterns across the 8×12 recording grid in each subject were analyzed using multivariate pattern analysis, whereby recordings from all recording contacts are provided to a classifier on a single-trial basis. The classifier was trained to discriminate the frequency of each pure-tone stimulus based on the brain responses recorded simultaneously from all contacts. The performance of the trained classifier in discriminating the frequency of pure-tone stimuli of unknown frequency provided an objective measure of the presence of stimulus-specific information in the aggregate brain response.

A sparse binomial logistic regression model was used as a base classification algorithm. Both *l1* and *l2* norm regularization were imposed on the weights during parameter estimation. As a result, this algorithm had advantages in the preservation of sparsity and correlated features. We used component-wise updating procedure (Shevade and Keerthi 2003; Cawley and Talbot 2005, 2006; Krishnapuram et al. 2005; Ryal et al. 2010). The classifier was applied to the data in a pairwise fashion (one vs. one scheme), and code words

were assigned to each class based on the classifier outputs using error-correcting output coding (Dietterich and Bakiri 1995). Decoding was done by computing Hamming distances (Allwein et al. 2001) between the coding matrix, $M=(-1, 0, 1)$, and the code words from the classifier’s output for each trial. The trials were assigned to the class which had the shortest Hamming distance to its corresponding row within the coding matrix, M .

The classification algorithm was applied for each time window to investigate the time course of the tone frequency discrimination pattern of the responses. Only contacts that had an average ERBP of at least 0.5 dB above baseline levels were selected for classification analysis. This thresholding procedure effectively reduced the number of features for the classification while avoiding potential over-fitting. Input to the classification algorithm was therefore a matrix of single-trial ERBP values for each contact and a time window, normalized in terms of its column mean and variance, and the vectors that assigned the class labels for each trial. The 2 regularization parameters (for *l1* and *l2* norm) were set by a grid in a common logarithmic scale to optimize its cross-validated classification accuracy.

Classification performance was assessed as a 5-fold cross-validated tone frequency identification accuracy. Average absolute values of weights adjusted for the accuracy of the final model (cross-validated) were calculated (overall weights; W) as a measure of the overall contribution for the discriminative pattern formation provided by each recording site:

$$W = \frac{K(K-1)}{2} \sum_{1 \leq i < j \leq K} A_{ij} |w_{ij}|$$

where A_{ij} is the classification accuracy between classes i and j , and K is the number of classes. Classification confusion matrices and pairwise classification results were also generated and presented.

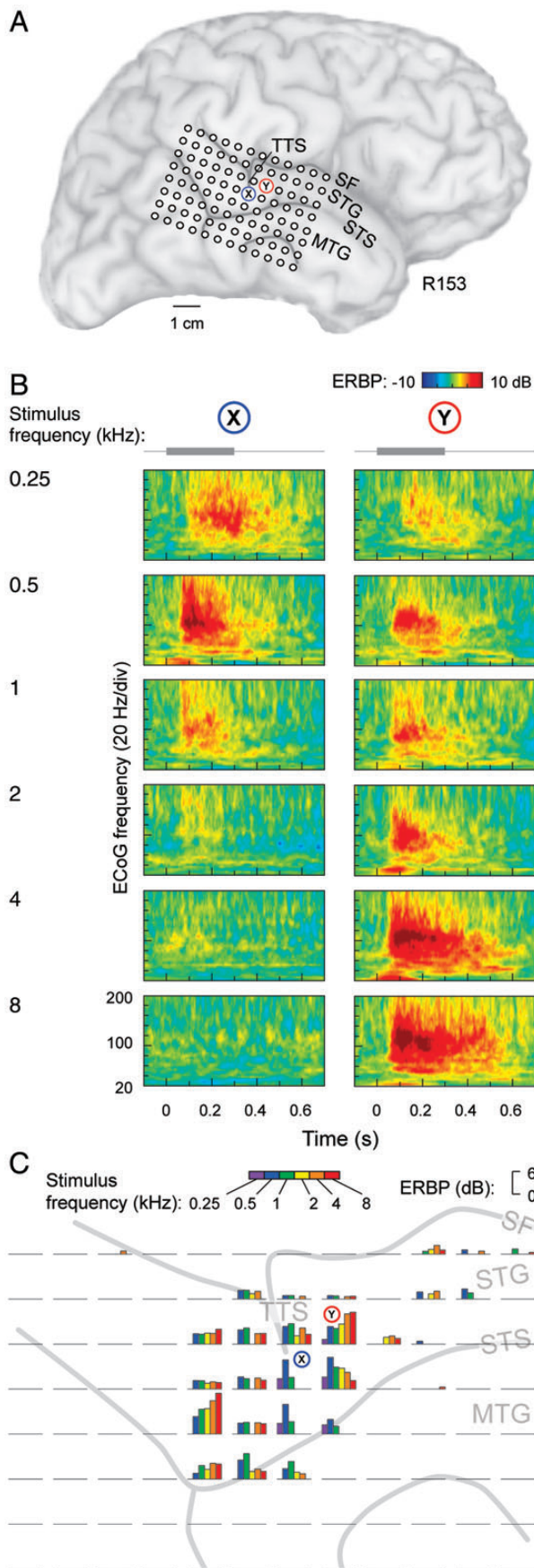
The effects of stimulus level on tone-elicited cortical activity were examined as frequency–intensity receptive fields that characterized individual recording sites. This was done by plotting high gamma ERBP values, averaged within 50-ms time intervals, as functions of stimulus frequency and intensity. In addition, classification analysis of single-trial responses to the 6 frequencies used in the main experimental set was done for data subsets containing responses to pairs of adjacent tone intensities and responses pooled across all intensities (except the lowest, 11 dB SPL).

Results

General Response Properties

Robust gamma-band responses were elicited in PLST in a tone frequency-specific manner. This enhanced activity, maximal in the high gamma range (75–150 Hz), is exemplified by tone-elicited responses recorded from a subdural grid in one of the subjects (Fig. 1). The presentation level in this case was 46 dB SPL. The recording grid covered PLST with extension onto the middle temporal gyrus (Fig. 1A). Examples of responses to low- and high-frequency tones (0.5 and 4 kHz, respectively) from 2 contacts on PLST are shown in Figure 1B. At contact X, the 0.5 kHz tone elicited the largest response, whereas at contact Y, the largest responses were elicited by the 2 highest frequencies (4 and 8 kHz). Typically, high gamma ERBP elicited by tone stimuli had onset latencies of about 50 ms and reached its maximum between 100 and 200 ms. This timing is similar to that observed for responses elicited by speech sounds (Steinschneider et al. 2011).

Tone-elicited responses recorded from lateral surface grids were confined almost exclusively to PLST and typically were observed in close proximity to the transverse temporal sulcus (TTS). In Figure 1C, ERBP measured in the high gamma



frequency range is depicted for all 96 recording contacts and color-coded for the tone frequency of stimulation. Only statistically significant increases above the baseline levels are plotted. In this example, ERBP was averaged over the time interval of 100–150 ms after stimulus onset. Maximal activity occurred within circumscribed areas of PLST located immediately adjacent to the TTS. Responses were also observed at a small number of recording sites located more anteriorly along the STG.

Individual sites typically exhibited broadly tuned frequency selectivity at suprathreshold stimulus levels with frequency response areas often spanning several octaves (Fig. 1C). This broad tuning seen at individual sites could be the result of low spatial specificity inherent to low impedance subdural contacts recording ECoG. However, marked differences in tuning occurred at the adjacent recording sites, indicating spatial resolution better than the 5-mm interelectrode distances.

Frequency–Intensity Receptive Fields

To determine if frequency tuning becomes more specific at lower intensities, frequency–intensity receptive fields across the entire recording grid in 5 subjects were examined. Clinically, feasible durations of experimental sessions placed a limit on the number of tone frequencies that could be presented and on the number of stimulus presentations for each frequency–intensity combination. For instance, tone frequencies above 8 kHz were not examined. It was reasoned that stimuli with frequencies above this range are less important for human communication (Gelfand 1998), and therefore of somewhat less interest than frequencies within the range of speech. These limits preclude quantitatively accurate estimates of tuning properties such as tuning width at threshold and at suprathreshold intensities. Instead, qualitatively accurate properties of tuning typically observed are described.

Magnitude and bandwidth of excitatory responses recorded from individual cortical sites typically increased with stimulus intensity. Figure 2 presents a representative data set obtained from 1 subject (R180; grid placement shown in Fig. 2A). Shown is the time interval (125–175 ms) determined by classification analysis as providing the peak performance for tone discrimination (see below). The increase in magnitude and bandwidth of excitatory responses can be observed in the frequency–intensity receptive fields that characterized 2 representative recording sites, X and Y (Fig. 2B).

At site X, maximal activation occurred at 4 kHz and was invariant to changes in stimulus intensity. However, as intensity increased, a progressive broadening of activation developed

Figure 1. Tone-elicited responses recorded from the lateral STG in a representative subject (R153). (A) Location of the 96-contact subdural electrode grid (TTS: transverse temporal sulcus; SF: Sylvian fissure; STG: superior temporal gyrus; STS: superior temporal sulcus; MTG: middle temporal gyrus). X and Y denote the locations of 2 sites whose detailed response properties are shown in panel B. (B) Time–frequency analysis of cortical activity elicited by pure-tone stimuli between 0.25 and 8 kHz (top to bottom), recorded from 2 sites on the STG (left and right panels). Stimulus schematic is shown on top. ERBP was calculated relative to baseline power measured in the reference period between 100 and 200 ms prior to stimulus onset. (C) High gamma ERBP in response to tone stimuli presented at different frequencies (color bars), simultaneously recorded from the 96-contact grid (individual plots). High gamma band (75–150 Hz) ERBP was averaged within the time window of 100–150 ms after stimulus onset.

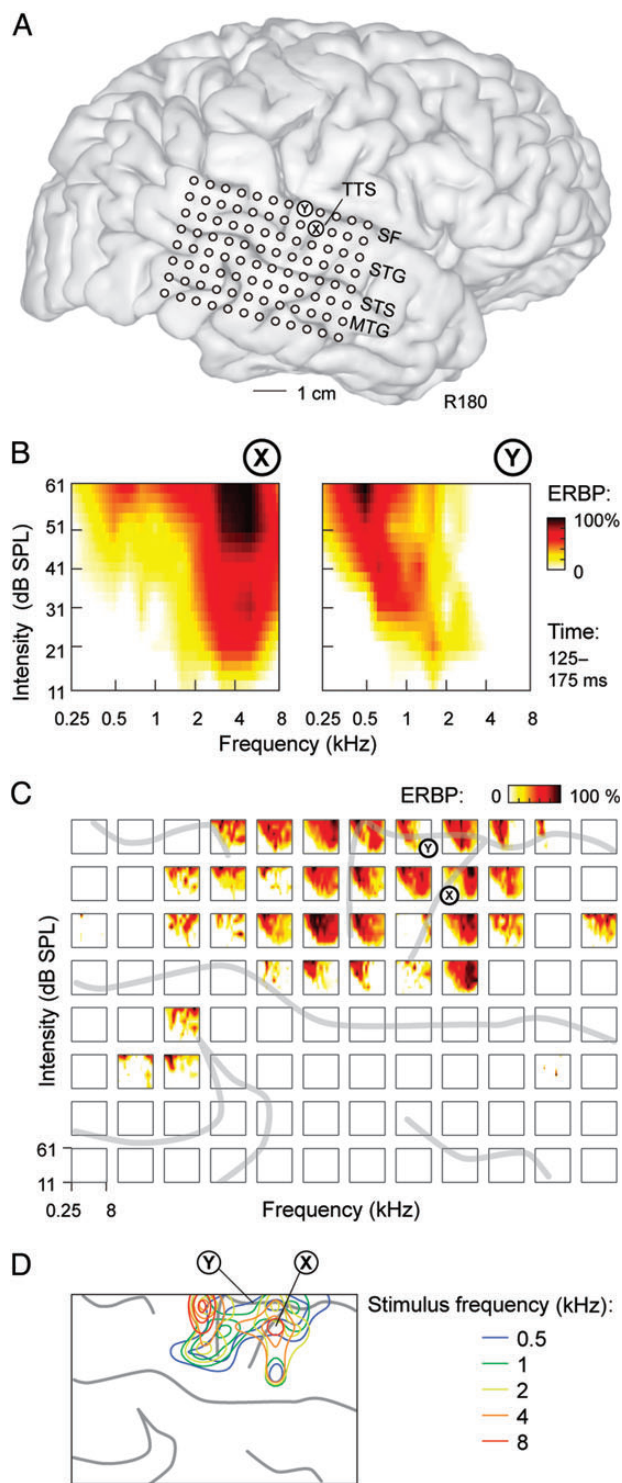


Figure 2. Frequency tuning properties of PLST in a representative subject (R180). (A) Location of the 96-contact subdural electrode grid (TTS: transverse temporal sulcus; SF: Sylvian fissure; STG: superior temporal gyrus; STS: superior temporal sulcus; MTG: middle temporal gyrus). X and Y denote the locations of 2 sites whose detailed response properties are shown in panel B. (B) Examples of frequency–intensity receptive fields for 2 sites in a representative subject. (C) Frequency–intensity receptive fields across the entire recording grid. Sulcal patterns are indicated by gray lines. (D) 50th, 75th, and 95th percentile contours of cortical responses to tones of 5 frequencies presented at 66 dB SPL. Sulcal patterns are shown by gray lines. Results shown in panels (B–D) are based on the analysis of ERBP averaged within 125–175 ms after stimulus onset.

for low frequencies. The absence of tone stimulation above 8 kHz precludes determining whether this response broadening would have occurred at higher tone frequencies. Regardless, this pattern was observed in multiple subjects and at many sites that were maximally responsive to high frequencies at low intensities.

A completely different frequency–intensity receptive field also commonly observed is exemplified by the response pattern shown at site Y. Here, responses at the lowest stimulus intensity were confined to the midrange frequencies (1.6 kHz) and also became more responsive to low-frequency tones at higher stimulus intensities. In contrast to the receptive field seen at the site X, maximal excitation was not invariant to intensity and instead was characterized by a downward shift in the tone frequencies that evoked the largest magnitude ERBP responses as stimulus intensity increased.

Other shapes of frequency–intensity receptive fields observed include those referred to as “mixed multip peaked” (Sutter 2000), a pattern characterized by broad and complex peaks of activation across frequencies and intensities. “Circumscribed” response patterns (Sutter 2000), characterized by tuning to selective frequencies at a selective range of submaximal stimulus intensities, were not observed. Overall, there was considerable variability in response patterns as a function of stimulus intensity (Fig. 2C), and this finding was characteristic of all 5 subjects, in whom frequency–intensity receptive fields were examined.

To ensure replicability of the activation patterns that characterized responses to pure-tone stimuli, response distributions elicited by a set of tones presented at 66 dB SPL in a separate experimental session were performed 2 days later (Fig. 2D). Consistent with the activation pattern obtained from the more extensive data set shown in Figure 2C, the region immediately posterior to the TTS remained maximally responsive to low frequencies, while ERBP preferentially elicited by high-frequency tones remained distributed anterior and posterior to this region.

Spatiotemporal Organization

Spectral specificity of PLST responses was temporally dynamic and was greatest at earlier portions of the neural responses. In these earlier time periods, an orderly spatial representation of tone frequency along PLST could often be observed (Fig. 3; subject R153, 50–100, 100–150 ms after stimulus onset). High gamma ERBP elicited by low-frequency stimuli was maximal immediately ventral to the TTS (Fig. 3A, left-hand columns). As tone frequency increased, maximal activity occurred at sites more anterior and posterior to those locations maximally excited by low-frequency tones. Higher tone frequencies failed to elicit significant responses at recording sites maximally activated by low frequencies, yielding mirror-image patterns of spectral specificity (Fig. 3B). An additional region maximally activated by low-frequency tones was observed at an even more anterior location, though limitations of grid coverage precluded more detailed mapping of this area.

While an orderly representation of tone frequency could be discerned in some cases (Figs 2D and 3B), spatial patterns exhibited significant variability across subjects. In the majority of subjects, clear mirror-image patterns were not observed,

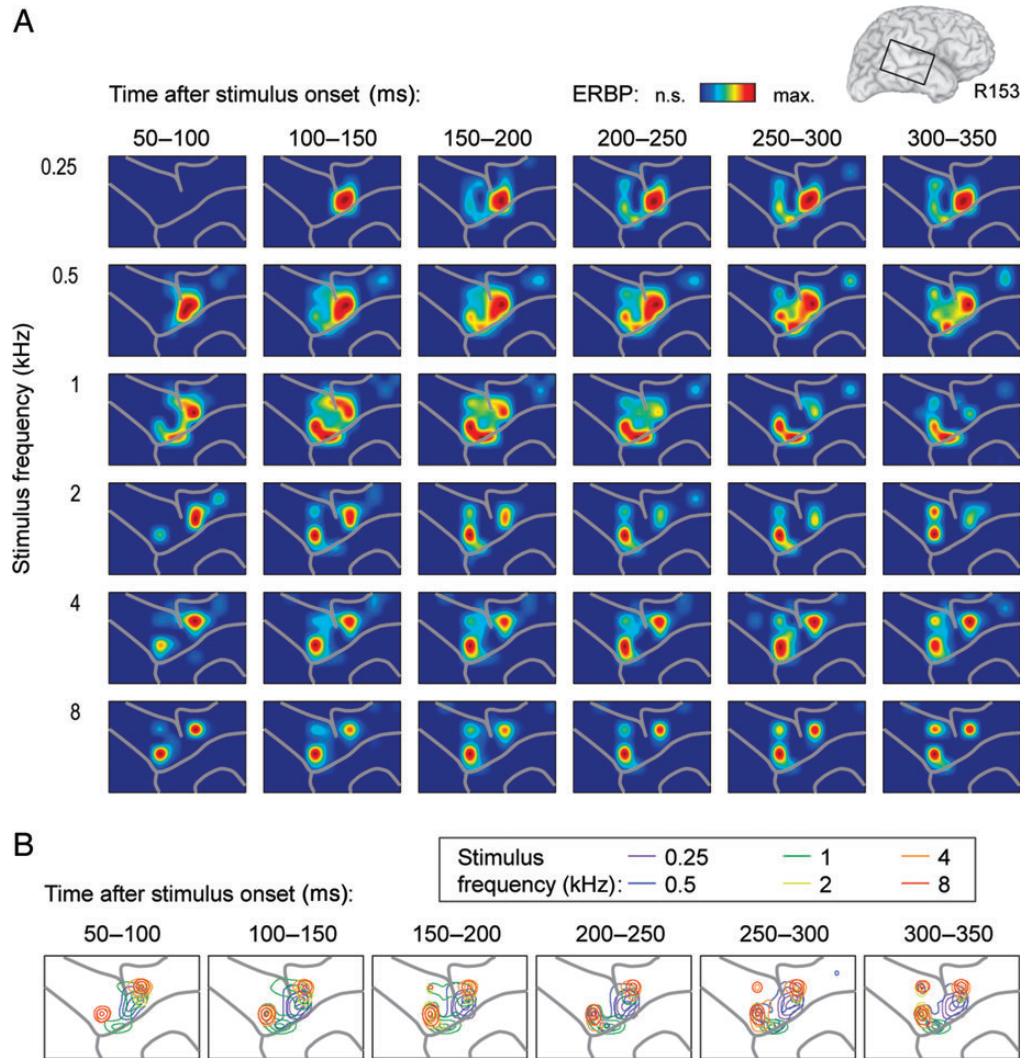


Figure 3. Spectral specificity of cortical high gamma responses with frequency gradients. Data recorded from the right hemisphere in a representative subject (R153). (A) Cortical activation maps for the 6 stimulus frequencies (top to bottom). The values of average high gamma ERBP within 6 consecutive 50-ms time windows (left to right) obtained for each contact were normalized relative to the maximum response across contacts for a given frequency and time interval and smoothed using cubic interpolation. Only significant responses (paired *t*-tests with false discovery rate correction; $q < 0.01$) are plotted. (B) Maps depicting 50th, 75th, and 95th percentile contours of cortical responses to the 6 stimulus frequencies. Sulcal patterns are shown by gray lines.

and more complex spatial distributions of frequency responses were noted instead (Fig. 4, subject L178). In this example, tones were presented at 66 dB SPL. Low frequencies preferentially activated multiple sites located anteriorly on the left hemisphere grid during early time intervals after stimulus onset, while higher frequency tones initially activated multiple sites distributed at more posterior locations (Fig. 4A, left-hand columns). Foci activated by both low- and high-frequency stimuli to a comparable degree were also observed (Fig. 4B).

In all cases, cortical activation patterns were temporally dynamic. As a general rule, sites initially responsive only to high-frequency tones become significantly activated by low tone frequencies during later time intervals (Figs 3A and 4A, right-hand columns). In Figure 3, activation in response to low-frequency tones was initially localized adjacent to the TTS, and then expanded into surrounding sites that were initially responsive to only high-frequency tones. In Figure 4, areas initially activated by high frequencies became progressively activated by low-frequency tones as well. This transition

in spectral specificity became especially prominent beyond 150 ms after stimulus onset. We did not observe the converse pattern, wherein sites initially selectively responsive to low-frequency tones would become responsive to high-frequency stimuli over time.

Classification Analysis

Spectral specificity within PLST was computationally evaluated across the entire data set by determining the degree to which neural activity could predict tone frequency. To this end, a sparse logistic regression-based classifier algorithm was applied to the single-trial high gamma ERBP data recorded from the entire grid in each subject (Fig. 5). Classification was performed on averaged power binned in 50-ms time intervals in sequential 25 ms steps (range -100 – 50 ms to 675725 ms relative to stimulus onset).

Classifier accuracy was above chance (16.7%) in all subjects. Figure 5A,B depicts classification accuracy for

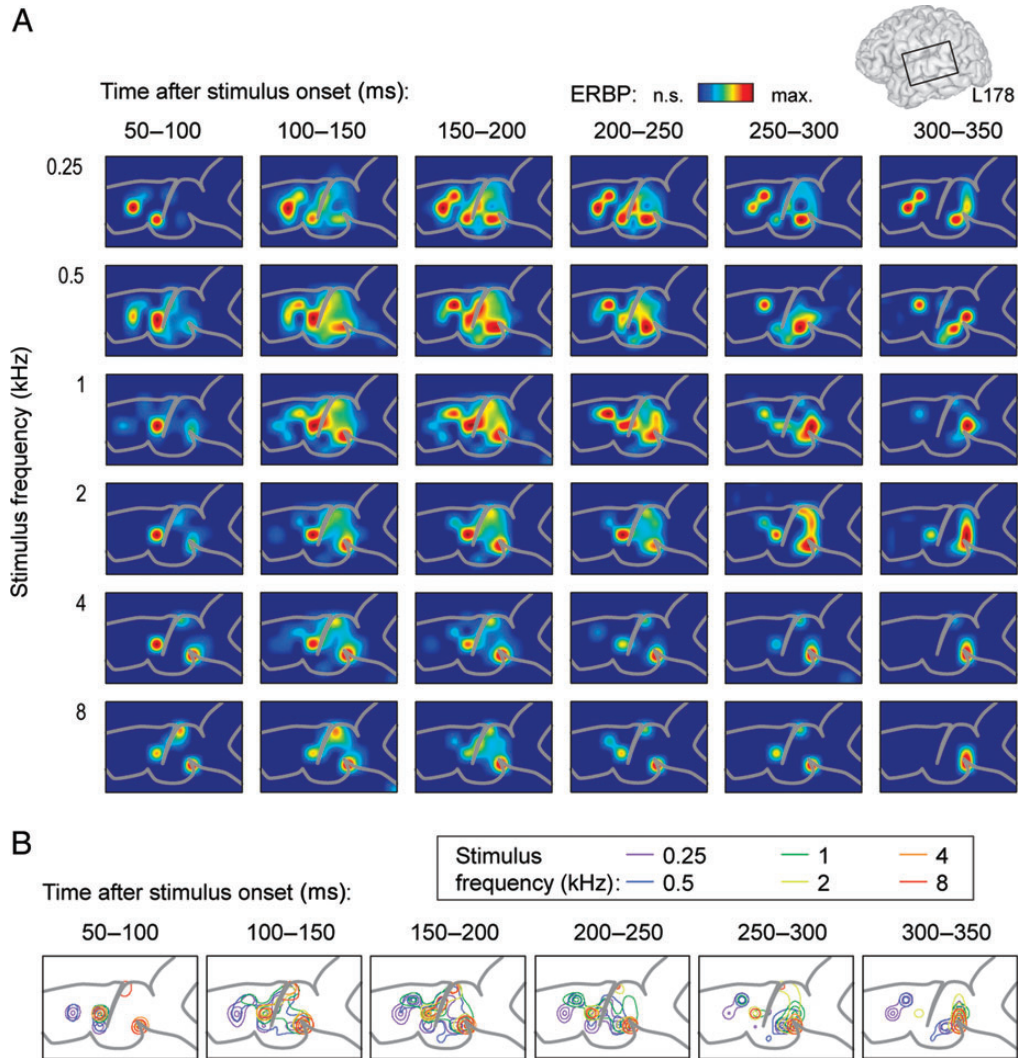


Figure 4. Spectral specificity of cortical high gamma responses without frequency gradients. Data recorded from the left hemisphere in a representative subject (L178). See legend of Figure 2 for details.

discrimination between the 6 tone stimuli as a function of time interval for the left-hemispheric and right-hemispheric subjects, respectively. Maximum accuracy (indicated by circles in Fig. 5A,B) could reach over 60% and occurred early in the responses in all subjects. It was most commonly observed in the 100–150-ms time interval, with a subsequent decay in performance. Both mirror-image and clustered organizational patterns (see examples for subjects R153 and L178 in Figs 3 and 4, respectively) could exhibit comparable accuracy in tone discrimination by classification analysis.

As left and right hemispheres have been proposed to differentially process temporal and spectral information (e.g. Zatorre et al. 2002; Zatorre and Gandour 2008), we compared peak accuracies between the 2 hemispheres. No significant difference was observed across hemispheres ($t(11) = -0.556$, $P = 0.589$ 2-tailed; Fig. 4C). Mean peak discrimination accuracy across all 13 subjects was 43.5% (95% confidence interval 37.0–50.1%) and was significantly greater than chance (1 sample $t(12) = 8.926$, $P < 0.000005$ 1-tailed). To ensure that overall trends were not driven by several subjects with the highest classification performance, the mean accuracy and

95% confidence intervals were calculated across subjects (Fig. 5D). Averaged across subjects, performance was above chance beginning at the 25–75-ms interval, peaked at 75–125 ms, and then slowly decayed to the chance levels at 575–625 ms after stimulus onset.

Classification errors were not randomly distributed and typically occurred between tones of adjacent frequencies. This is demonstrated by confusion and pairwise classification matrices for an exemplary subject (R153; Fig. 6, left column) and for the entire data set (Fig. 6, right column). When stimuli were misclassified, predicted stimuli were typically the next nearest neighbors in frequency to the presented tone (Fig. 6A). Likewise, pairwise classification results demonstrate greatest accuracy between those stimuli that were progressively more dissimilar in frequency (Fig. 6B). Thus, these findings are consistent with spectrally specific activation patterns.

Classification Accuracy Across Stimulus Intensities

Results indicate that ERBP recorded from PLST contains sufficient information to discriminate tone frequencies

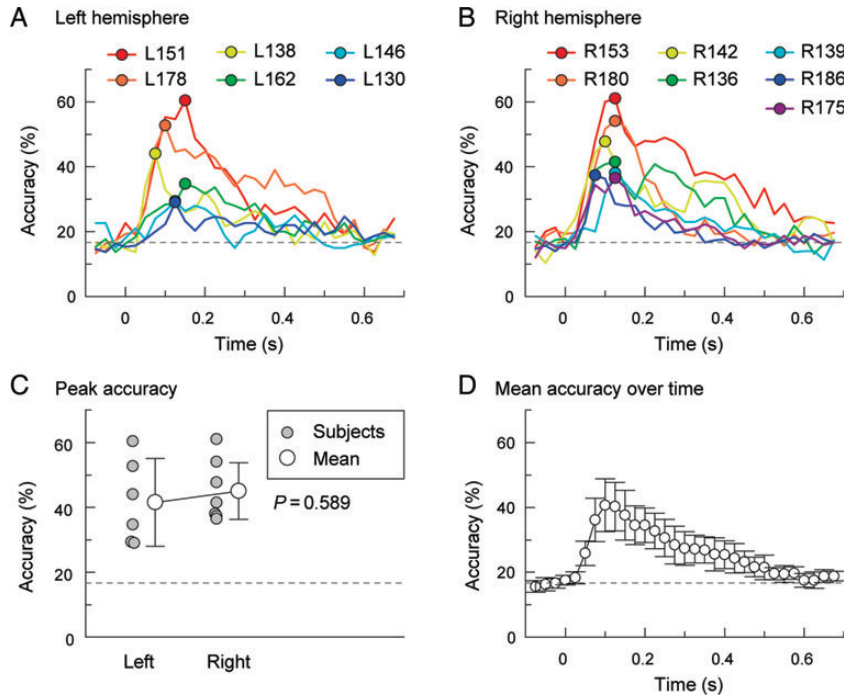


Figure 5. Tone classification analysis based on cortical high gamma ERBP. (A) Classification accuracy for the 6 pure-tone stimuli (0.25, 0.5, 1, 2, 4, and 8 kHz) as a function of time window in the 6 subjects implanted in the left hemisphere. Circles indicate maximum classification accuracy in each subject. (B) Classification accuracy for the 6 pure-tone stimuli as a function of time window in the 7 subjects implanted in the right hemisphere. (C) Maximum classification accuracy for recordings from left- and right-hemisphere cases. Gray circles correspond to performance maxima in individual subjects; white circles are mean values, error bars correspond to 95% confidence intervals. (D) Classification accuracy over time. Circles indicate mean values; error bars represent 95% confidence intervals. Dashed lines indicate chance performance (16.7% for 6 frequencies).

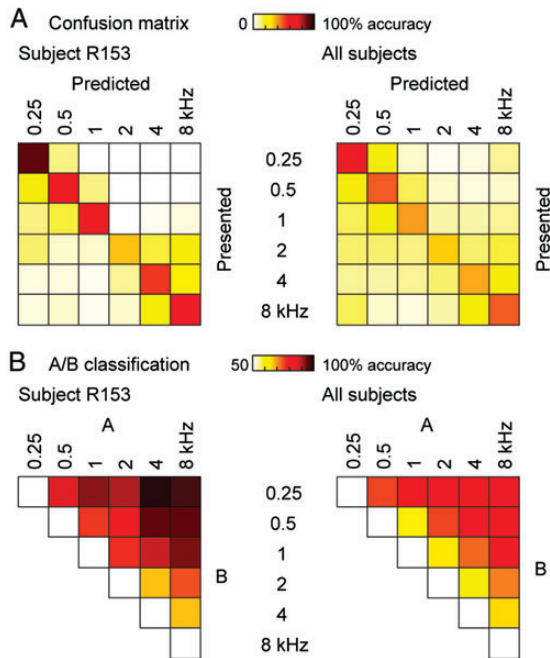


Figure 6. Classification accuracy based on the data from a representative subject (R153), and data averaged across the 13 subjects (left and right panels, respectively), measured in time intervals of the highest overall classification performance (indicated by circles in Fig. 4A,B). (A) Classification confusion matrices. (B) A/B pairwise classification.

significantly above chance level. A key question is whether this accuracy is invariant to changes in stimulus intensity. Increases in the tone level modulate frequency selectivity

(Fig. 2) and, as shown with fMRI, lead to an enhanced cortical activation (Woods et al. 2010). This variation in activation based on stimulus intensity may confound physiological classification analysis of tone frequency.

To examine the effects of stimulus intensity on tone discrimination, we evaluated classifier accuracy based on responses to tones presented at different intensities within 10 dB ranges (Fig. 7). For subject R180, whose frequency–intensity receptive fields are shown in Figure 2, classifier accuracy peaked at about 55–60% and, similar to earlier classification analyses, was maximal early in the response (Fig. 7, top left panel). Classification accuracy and its time course were similar when responses to tones presented at different intensities were examined (color lines in Fig. 7).

Similarity in classification performance at different stimulus intensities does not, however, conclusively demonstrate that activity within PLST contains sufficient information to decode tone frequency. That is, because it is possible that different response features are used for the representation of frequency at multiple intensities. A more powerful demonstration of classification based solely on stimulus frequency would be to train the classifier on neural responses randomly sampled from trials of all intensities (21–61 dB SPL), and test classifier performance on trials also randomly selected across intensities. Therefore, the classifier is blinded to stimulus intensity and can only be making discriminations based on stimulus frequency.

Importantly, when the classifier was trained and tested on tones culled from random intensity trials, overall accuracy remained essentially unchanged (black line in Fig. 7, top left panel). This finding indicates that PLST contains enough information to discriminate pure-tone frequencies regardless

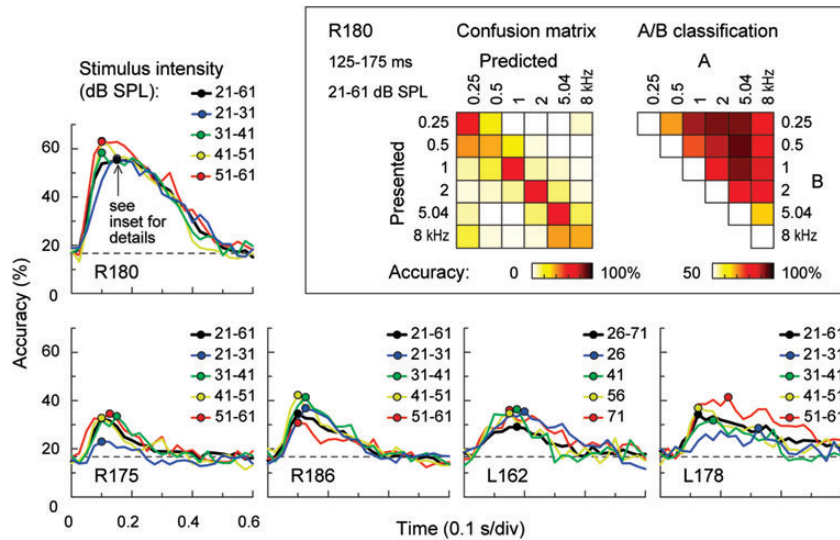


Figure 7. Frequency discrimination accuracy across stimulus intensities. Classification accuracy for the 6 pure-tone stimuli (0.25, 0.5, 1, 2, 5.04, and 8 kHz) is plotted as a function of time window for trials presented at different intensities. Circles indicate maximum classification accuracy for each intensity condition. Data from 5 subjects are shown. Inset: Classification confusion matrix (left) and pairwise classification accuracy (right) for stimulus intensities between 21 and 61 dB SPL in subject R180 for the time interval of peak classification performance (75–125 ms after stimulus onset).

of intensity. Further, detailed examination showed that recording sites that contributed most to the performance of the classifier in this test (as determined by classifier weights) were typically those that demonstrated maximal excitation by a specific frequency across stimulus intensities (data not shown).

Both confusion and pairwise classification matrices continued to show that errors were not random and were most likely to occur to the next nearest neighbor in frequency (Fig. 7, inset). While performance was not as optimal as that shown for subject R180 in other subjects examined using this paradigm (Fig. 7, bottom row), they all showed above-chance performance when tone frequencies were collapsed across intensities ($P < 0.05$, 1-sided Wilcoxon signed-rank test). Mean peak discrimination accuracy was 37.1%, with a 95% confidence interval of 24.3–50.0% (chance = 16.7%). Responses exhibited the greatest specificity in the early time intervals following tone onset. Averaged across the 5 subjects over time, classifier performance peaked at 100–150 ms interval. Thus, although the spectral sensitivity of responses recorded from individual brain sites varied as a function of stimulus intensity, there is sufficient information in the responses distributed across PLST to enable the classifier to identify spectral information in a manner that is largely insensitive to differences in stimulus intensity.

Discussion

Summary of Findings

This study demonstrates that auditory cortex overlying the PLST is strongly activated by pure-tone stimuli and that activation patterns vary as a function of stimulus frequency. This activation occurs along a restricted portion of PLST, adjacent to the TTS. While responses at individual sites exhibit frequency selectivity, they are typically broadly tuned at supra-threshold levels. In several subjects, spatial organization of

frequency selectivity demonstrated a tonotopic pattern with mirror-image gradients centered on a low-frequency responsive region. As a rule, however, spectral organization is represented by clustered response patterns, wherein low- and high-frequency tones maximally activate different sites without clear spatially patterned frequency gradients. Regardless of the specific spatial pattern, classification analysis consistently yields above-chance performance, indicating that activity along PLST contains information that is sufficient to differentiate the pure tones that were presented in these experiments. Errors in classifier performance are most commonly biased toward a next nearest neighbor to the presented tone. Maximum accuracy is typically achieved at around 100–150 ms after stimulus onset and gradually declines. The broadening of spectral specificity over time occurs as regions initially responsive to a narrow range of frequencies become activated by a broader range of tones. This feature is particularly prominent for sites initially selectively activated by high-frequency tones. Classifier performance is largely unaffected by changes in stimulus intensity over a 40 dB range, and we observed no significant difference between classification results in left-sided and right-sided PLST recordings.

Tuning Characteristics of PLST

A major advantage of examining high gamma activity in the ECoG is its positive correlations with both spiking activity and hemodynamic responses (Nir et al. 2007; Steinschneider et al. 2008; Whittingstall and Logothetis 2009). This enables high gamma activity measurements to serve as a bridge between different research techniques and facilitates comparisons across studies performed using different methodological approaches. Although it is difficult to extrapolate tuning characteristics as measured by high gamma ECoG activity to those obtained from single units, the broad tuning on PLST is in contrast to sharper tuning observed in the primary auditory cortex of humans estimated from single-unit recordings (Howard et al. 1996; Bitterman et al. 2008). Our results also

differ from the sharp tuning identified using similar ECoG measures in the monkey primary auditory cortex (Brosch et al. 2002; Steinschneider et al. 2008). On the other hand, the current finding is in concordance with more complex and broadly tuned responses observed in lateral nonprimary areas using similar ECoG measures in humans (Pasley et al. 2012) and single-unit responses in nonhuman primates (Recanzone et al. 2000; Rauschecker and Tian 2004).

Common to PLST were tuning curves characterized by responses restricted to higher frequencies at low intensities and a broadening of excitation at higher intensities (e.g. Fig. 2, site X). This property is typical of many tuning curves seen in the primary auditory cortex of nonhuman primates and cats (e.g. Recanzone et al. 2000; Sutter 2000). Future work with frequencies beyond 8 kHz will be required to more accurately characterize the shape of frequency–intensity receptive fields and to determine whether the broadening of excitation is symmetric or asymmetric around the best frequency.

An additional common shape observed at sites on PLST was characterized by a shift in tuning specificity toward lower frequencies at higher intensities (e.g. Fig. 2, site Y). This shape resembles the “slant-lower” tuning curve profile described in single-unit recordings in A1 in the anesthetized cat (Sutter 2000). Similar patterns have been reported from core fields A1 and R, and caudomedial belt field CM in the awake macaque (cf. Figs 4 and 7 in Recanzone et al. 2000). Another frequency–intensity receptive field shape commonly observed was characterized by multiple peaks of spectral sensitivity at both lower and higher intensities. This mixed multi-peaked pattern is uncommon in A1 (Sutter 2000). Our protocol was not designed to test whether multiple peaks observed in these cases were harmonically related based upon limited number of tone frequencies tested. Finally, the circumscribed pattern, commonly observed in the primate primary auditory cortex (Sadagopan and Wang 2010), was not seen in our data set.

Relationship to Functional Neuroimaging Studies

Previous attempts to characterize the spectral response properties of PLST using fMRI have led to contradictory results. Our findings are incompatible with the conclusion that PLST does not respond differentially to pure tones of different frequencies (Formisano et al. 2003; Lewis et al. 2009). In other studies, the spectral organization of PLST was examined as a component of a larger analysis focusing on frequency selectivity within the superior temporal plane (Talavage et al. 2004; Humphries et al. 2010). Both studies observed a small anterior-to-posterior gradient of frequency sensitivity from high to low frequencies, respectively. These observations may reflect a portion of the mirror-image gradients occasionally observed in the present study. Another study (Woods et al. 2010) has identified frequency selectivity on PLST that might represent a functional imaging analog of the more general organizational principles identified in the current electrophysiological study; individual sites were frequency selective but there was no obvious spatial pattern correlating frequency selectivity with location along PLST. Our findings did not confirm the dorsal-to-ventral spatial orientation of tonotopic patterns on the lateral STG reported by Striem-Amit et al. (2011) using fMRI mapping techniques.

The present study using direct intracranial recording methods identified several organizational features that should be considered when interpreting the findings from fMRI studies.

First, the manner in which spectral information is represented in the brain responses recorded from different PLST sites changes rapidly over time following stimulus onset. Existing fMRI techniques cannot resolve the changing activation patterns with this degree of temporal resolution, and the likely result is a spectral “blurring” of blood oxygen level-dependent (BOLD) signal responses. The BOLD signal therefore may not capture the spectral sensitivity of the earliest portion of the neural response. Secondly, we did not identify a single spatially organized, spectrally specific, response pattern that could characterize all subjects. This would confound the interpretation of fMRI studies trying to identify a unified organizational scheme. Thirdly, variations in tone selectivity based on the stimulus level (particularly at the loud levels typically utilized in fMRI paradigms), stimulation paradigm (e.g. pure-tone vs. frequency sweep stimuli, continuous vs. sparse acquisition), and background environment can potentially obscure underlying organization. All these factors may have contributed to the differences between the current data and those obtained using fMRI.

Relationship to Multifield Models of Auditory Cortex Organization

One purpose of the current experiments was to provide additional insights into what role the PLST cortex may play in multifield models of human auditory cortex functional organization. These models reflect concepts derived from extensive experimental work carried out in nonhuman primates that demonstrate the existence of multiple hierarchically organized auditory fields distributed within core, belt, and parabelt cortical regions (Kaas and Hackett 1998). Cytoarchitectonic studies of the human auditory cortex offer different interpretations for the place of PLST within the hierarchy of auditory cortical fields. Some studies interpret it as a belt area (Galaburda and Sanides 1980; Fullerton and Pandya 2007), others (e.g. Rivier and Clarke 1997) refer to it as a “downstream” association cortex (i.e. parabelt), while Sweet et al. (2005) do not include it in the core–belt–parabelt model altogether. An analysis of the anatomical studies of human auditory cortex concluded that no clear designation of PLST could be made based on current evidence (Hackett 2007).

In the setting of experimental animal research, individual fields are defined by the combinations of field-specific features, including gross anatomical location, cytoarchitectonics, spatial orientation relative to other fields, connectivity patterns, and functional properties of the populations of neurons distributed throughout a field. Limitations inherent to human subject research preclude using the exact same powerful experimental methods to precisely define and compare auditory field homologies across primate species. It is possible, however, to make certain inferences based on experimental evidence that can be safely collected in humans. In reference to PLST, we know from earlier work that the functional properties of this cortical region differ markedly from the properties of core auditory cortex located on posterior-medial HG (Howard et al. 2000; Brugge et al. 2008). These studies lead to the conclusion that area PLST is comprised of noncore

auditory cortex positioned lateral to the known core cortex. Because grid recording arrays are not placed on the superior temporal plane within the Sylvian fissure, it is not technically feasible to create fine spatial grain maps of the expanse of planum temporale cortex located between known core cortex (posteromedial HG) and PLST. This precludes directly mapping any boundaries that may exist within the planum temporale demarcating transitions between core, belt, and possibly parabelt fields.

Another criterion—functional connectivity between auditory fields—cannot be directly extrapolated from experimental animal studies that were performed using histological tract tracing methods. In human subjects, electrical stimulation tract tracing methods have been used to show that short latency functional connections exist between posterior-medial HG (core cortex) and sites within PLST (Brugge et al. 2003). However, this method is not capable of resolving whether the observed functional connection represents the activation of mono- or polysynaptic pathways.

The current finding in some subjects of a mirror-image tonotopic pattern of frequency representation within PLST in response to pure-tone stimuli is similar to the patterns observed in the nonhuman primate lateral belt cortex (Rauschecker and Tian 2004). For the reasons described above, there is insufficient experimental information to definitively assign field homology. However, the current observations of robust, spectrally selective responses to pure-tone stimuli within a circumscribed region of the lateral cortex, combined with earlier findings of short latency functional connections with core cortex are consistent with PLST occupying a relatively early stage in hierarchical models of the multifield processing of auditory information.

Ongoing studies address these issues by combining preoperative fMRI mapping with direct intracranial recordings in the same subjects. By determining relationships between BOLD signal changes and direct electrophysiological recordings, and making use of the unrestricted spatial sampling capacity of MRI techniques, a combined approach has the potential to enhance our ability to define human auditory cortex fields and their boundaries (Chevillet et al. 2011). The presence and topographic patterns of pure-tone responses recorded from PLST will be a key physiological feature, facilitating the reliable identification of specific auditory fields in individual subjects.

Spatiotemporal Representation of Frequency

The cortical spatial patterns of frequency representation within PLST change as a function of the time period examined following stimulus onset. Recordings from individual brain sites show that the greatest degree of spectral selectivity is displayed during the initial portions of the response. Later in the response, stimuli with a broader range of frequencies evoke increases in gamma-band power. This pattern of changing spectral specificity within cortical sites over time may reflect integrative cortical processes that are also engaged when humans hear more spectrally complex acoustic stimuli, such as vowels containing multiple formants. Additional experimental studies designed specifically to investigate this integrative process, and its relevance to speech sound processing, are underway.

Psychophysical studies demonstrate that, as tone intensity decreases, difference limens increase (Wier et al. 1977). From these perceptual studies, one might expect concordant physiological changes in brain activation patterns. We found that classification performance was stable over a 40-dB range of stimulus intensities. However, difference limens are much smaller than the spectral separations of the tones used in this study. Therefore, we would be unable to identify the underlying physiological correlates of intensity-related changes in difference limens with the paradigm used in this study.

Left- Versus Right-Hemispheric Processing

There was no significant difference in classifier performance between left- and right-sided PLST subjects. This suggests that high gamma activity evoked within PLST in the 2 hemispheres contains the comparable levels of information about the frequency of the pure-tone stimuli that were delivered. This finding is compatible with current models of auditory cortical processing positing that PLST on both hemispheres is involved in acoustic-to-phonetic transformations (Poeppel et al. 2008; Hickok 2009). This model stipulates that both hemispheres support accurate representation of the spectrotemporal attributes of the acoustic signal (Pasley et al. 2012). The bilateral representation of tone frequency information within PLST is consistent with this model. It has also been hypothesized that nonprimary auditory cortex on the left hemisphere plays a specialized role in the processing of acoustic stimulus features contained within shorter time epochs (e.g. speech formant transitions), whereas the same regions on the right hemisphere are preferentially involved in the processing of slower spectral features (e.g. those characterizing vowels and speech prosody; Poeppel 2001; Zatorre and Belin 2001; Zatorre et al. 2002; Boemio et al. 2005; Zatorre and Gandour 2008). The acoustic stimuli used in the current experiments were not designed to examine the type of hemispheric asymmetry in temporal sampling posited by this theory. Future work using electrophysiological measures will be required to assess whether differences in hemispheric processing at the level of PLST can be identified.

Intersubject Variability

Data obtained from individual subjects exhibited variability in spatial patterns of tone-elicited responses and in results obtained from classification analysis. There are many potential explanations for this variability, including differences in cortical anatomy across subjects and the locations of electrodes in individual subjects. For example, anatomically defined PLST is in direct continuity with the cortex of the superior temporal plane, a brain region from which recordings were not obtained. It is possible that a specific nonprimary auditory cortical field that is present on PLST in some subjects is confined to the immediately adjacent superior temporal cortex in other subjects. In this case, if an auditory field is tonotopically organized, surface recording arrays would only detect the tonotopic pattern in subjects with the field located on PLST.

Classification accuracy of pure-tone stimuli varied widely across subjects, with maximum performance ranging from about 30% to 60%. Another potential source of variability was the type of classification analysis employed in the study. Here, we made use of high gamma ERBP values measured within discrete time bins. Future studies using alternative

approaches, such as methods incorporating temporally distributed high gamma activity (e.g. Pasley et al. 2012) or lower frequency evoked components of the ECoG (e.g. Chang et al. 2010), may enhance accurate decoding of stimulus frequency and provide additional insights into the functional organization of PLST.

Another important consideration is subject attention, which is known to influence cortical processing within nonprimary auditory cortex (Mesgarani and Chang 2012). This variable is not optimally controlled during passive listening, and experiments specifically designed to address this issue are ongoing. Also, the current experiments were performed in neurosurgical patients with medically intractable epilepsy. Auditory cortex may be dysfunctional to varying degrees in these patients (Boatman and Miglioretti 2005), and antiepileptic medications may also affect the brain functions being studied. Although there are many possible etiologies for the observed intersubject variability, these would not explain the strongly positive findings that we consistently observed across subjects. Finally, although no significant correlation was found between classification accuracy and verbal comprehension index scores, the variability of spatial response patterns and classification accuracy may also be related to differences in auditory processing skills that we did not test.

Funding

This work was supported by the NIH (grant numbers R01-DC04290, R01-DC00657, UL1RR024979), Hearing Health Foundation (Collette Ramsey Baker Award), and the Hoover Fund.

Notes

We thank John Brugge, Richard Reale, Olaf Kaufman, Christopher Kovach, Haiming Chen, and Rachel Gold for help with experiment design, data collection, and analysis. *Conflict of Interest*: None declared.

References

- Aitkin LM, Webster WR. 1971. Tonotopic organization in the medial geniculate body of the cat. *Brain Res.* 26:402–405.
- Allwein E, Schapire R, Singer Y. 2001. Reducing multiclass to binary: a unifying approach for margin classifiers. *J Mach Learn Res.* 1:113–141.
- Benjamini Y, Drai D, Elmer G, Kafkari N, Golani I. 2001. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res.* 125:279–284.
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B.* 57:289–300.
- Besle J, Schevon CA, Mehta AD, Lakatos P, Goodman RR, McKhann GM, Emerson RG, Schroeder CE. 2011. Tuning of the human neocortex to the temporal dynamics of attended events. *J Neurosci.* 31:3176–3185.
- Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I. 2008. Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. *Nature.* 451:197–201.
- Boatman DF, Miglioretti DL. 2005. Cortical sites critical for speech discrimination in normal and impaired listeners. *J Neurosci.* 25:5475–5480.
- Boemio A, Fromm S, Braun A, Poeppel D. 2005. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat Neurosci.* 8:389–395.
- Brosch M, Bodinger E, Scheich H. 2002. Stimulus-related gamma oscillations in primate auditory cortex. *J Neurophysiol.* 87:2715–2725.
- Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA III. 2009. Coding of repetitive transients by auditory cortex on Heschl's gyrus. *J Neurophysiol.* 102:2358–2374.
- Brugge JF, Volkov IO, Garell PC, Reale RA, Howard MA III. 2003. Functional connections between auditory cortex on Heschl's gyrus and on the lateral superior temporal gyrus in humans. *J Neurophysiol.* 90:3750–3763.
- Brugge JF, Volkov IO, Oya H, Kawasaki H, Reale RA, Fenoy A, Steinschneider M, Howard MA III. 2008. Functional localization of auditory cortical fields of human: click-train stimulation. *Hear Res.* 238:12–24.
- Cawley GC, Talbot NLC. 2005. The evidence framework applied to sparse neural logistic regression. *Neurocomputing.* 64:119–135.
- Cawley GC, Talbot NL. 2006. Gene selection in cancer classification using sparse logistic regression with Bayesian regularization. *Bioinformatics.* 22:2348–2355.
- Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT. 2010. Categorical speech representation in human superior temporal gyrus. *Nat Neurosci.* 13:1428–1432.
- Chevillet M, Riesenhuber M, Rauschecker JP. 2011. Functional correlates of the anterolateral processing hierarchy in human auditory cortex. *J Neurosci.* 31:9345–9352.
- Crone NE, Boatman D, Gordon B, Hao L. 2001. Induced electrocorticographic gamma activity during auditory perception. *Clin Neurophysiol.* 112:565–582.
- Da Costa S, van der Zwaag W, Marques JP, Frackowiak RS, Clarke S, Saenz M. 2011. Human primary auditory cortex follows the shape of Heschl's gyrus. *J Neurosci.* 31:14067–14075.
- Dietterich TJ, Bakiri G. 1995. Solving multiclass learning problems via error-correcting output codes. *J Artif Intell Res.* 2:263–286.
- Edwards E, Soltani M, Deouell LY, Berger MS, Knight RT. 2005. High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. *J Neurophysiol.* 94:4269–4280.
- Edwards E, Soltani M, Kim W, Dalal SS, Nagarajan SS, Berger MS, Knight RT. 2009. Comparison of time–frequency responses and the event-related potential to auditory speech stimuli in human cortex. *J Neurophysiol.* 102:377–386.
- Formisano E, Kim DS, Di Salle F, van de Moortele PF, Ugurbil K, Goebel R. 2003. Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron.* 40:859–869.
- Fullerton BC, Pandya DN. 2007. Architectonic analysis of the auditory-related areas of the superior temporal region in human brain. *J Comp Neurol.* 504:470–498.
- Galaburda A, Sanides F. 1980. Cytoarchitectonic organization of the human auditory cortex. *J Comp Neurol.* 190:597–610.
- Gelfand SA. 1998. Hearing: an introduction to psychological and physiological acoustics. 3rd ed. New York (NY): Marcel Dekker.
- Goll JC, Crutch SJ, Warren JD. 2010. Central auditory disorders: toward a neuropsychology of auditory objects. *Curr Opin Neurol.* 23:617–627.
- Gomez-Ramirez M, Kelly SP, Molholm S, Sehatpour P, Schwartz TH, Foxe JJ. 2011. Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: a human electrocorticographic investigation. *J Neurosci.* 31:18556–18567.
- Hackett TA. 2007. Organization and correspondence of the auditory cortex of humans and nonhuman primates. In: Kaas JH, editor. *Evolution of the nervous system.* Oxford, UK: Elsevier. p. 109–119.
- Hackett TA, Preuss TM, Kaas JH. 2001. Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J Comp Neurol.* 441:197–222.
- Hickok G. 2009. The functional neuroanatomy of language. *Phys Life Rev.* 6:121–143.
- Hickok G, Poeppel D. 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition.* 92:67–99.
- Howard MA, Nourski KV, Brugge JF. 2012. Invasive research methods. In: Fay RR, Popper AN, editors. *Springer handbook of auditory research—human auditory cortex.* New York (NY): Springer Science. p. 39–67.
- Howard MA III, Volkov IO, Abbas PJ, Damasio H, Ollendieck MC, Granner MA. 1996. A chronic microelectrode investigation of the

- tonotopic organization of human auditory cortex. *Brain Res.* 724:260–264.
- Howard MA, Volkov IO, Mirsky R, Garell PC, Noh MD, Granner M, Damasio H, Steinschneider M, Reale RA, Hind JE *et al.* 2000. Auditory cortex on the human posterior superior temporal gyrus. *J Comp Neurol.* 416:79–92.
- Humphries C, Liebenthal E, Binder JR. 2010. Tonotopic organization of human auditory cortex. *Neuroimage.* 50:1202–1211.
- Kaas JH, Hackett TA. 1998. Subdivisions of auditory cortex and levels of processing in primates. *Audiol Neurootol.* 3:73–85.
- Kaas JH, Hackett TA, Tramo MJ. 1999. Auditory processing in primate cerebral cortex. *Curr Opin Neurobiol.* 9:164–170.
- Kashino M, Kondo HM. 2012. Functional brain networks underlying perceptual switching: auditory streaming and verbal transformations. *Philos Trans R Soc Lond B Biol Sci.* 367:977–987.
- Krishnapuram B, Carin L, Figueiredo MA, Hartemink AJ. 2005. Sparse multinomial logistic regression: fast algorithms and generalization bounds. *IEEE Trans Pattern Anal Mach Intell.* 27:957–968.
- Langers DR, van Dijk P. 2012. Mapping the tonotopic organization in human auditory cortex with minimally salient acoustic stimulation. *Cereb Cortex.* 22:2024–2038.
- Leaver AM, Rauschecker JP. 2010. Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. *J Neurosci.* 30:7604–7612.
- Lewis JW, Talkington WJ, Walker NA, Spirou GA, Jajosky A, Frum C, Brefczynski-Lewis JA. 2009. Human cortical organization for processing vocalizations indicates representation of harmonic structure as a signal attribute. *J Neurosci.* 29:2283–2296.
- Merzenich MM, Brugge JF. 1973. Representation of the cochlear partition of the superior temporal plane of the macaque monkey. *Brain Res.* 50:275–296.
- Mesgarani N, Chang EF. 2012. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature.* 485:233–236.
- Morel A, Garraghty PE, Kaas JH. 1993. Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. *J Comp Neurol.* 335:437–459.
- Nir Y, Fisch L, Mukamel R, Gelbard-Sagiv H, Arieli A, Fried I, Malach R. 2007. Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. *Curr Biol.* 17:1275–1285.
- Obleser J, Zimmermann J, Van Meter J, Rauschecker JP. 2007. Multiple stages of auditory speech perception reflected in event-related fMRI. *Cereb Cortex.* 17:2251–2257.
- Oya H, Kawasaki H, Howard MA III, Adolphs R. 2002. Electrophysiological responses in the human amygdala discriminate emotion categories of complex visual stimuli. *J Neurosci.* 22:9502–9512.
- Pasley BN, David SV, Mesgarani N, Flinker A, Shamma SA, Crone NE, Knight RT, Chang EF. 2012. Reconstructing speech from human auditory cortex. *PLoS Biol.* 10:e1001251.
- Poeppl D. 2001. Pure word deafness and the bilateral processing of the speech code. *Cogn Sci.* 25:679–693.
- Poeppl D, Idsardi WJ, van Wassenhove V. 2008. Speech perception at the interface of neurobiology and linguistics. *Philos Trans R Soc Lond B Biol Sci.* 363:1071–1086.
- Rauschecker JP, Tian B. 2004. Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol.* 91:2578–2589.
- Reale RA, Calvert GA, Thesen T, Jenison RL, Kawasaki H, Oya H, Howard MA, Brugge JF. 2007. Auditory-visual processing represented in the human superior temporal gyrus. *Neuroscience.* 145:162–184.
- Recanzone GH, Guard DC, Phan ML. 2000. Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *J Neurophysiol.* 83:2315–2331.
- Rivier F, Clarke S. 1997. Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. *Neuroimage.* 6:288–304.
- Romanski LM, Averbeck BB. 2009. The primate cortical auditory system and neural representation of conspecific vocalizations. *Annu Rev Neurosci.* 32:315–346.
- Rose JE, Greenwood DD, Goldberg JM, Hind JE. 1963. Some discharge characteristics of single neurons in the inferior colliculus of the cat I: tonotopic organization, relation of spike-counts to tone intensity and firing patterns of single elements. *J Neurophysiol.* 26:294–320.
- Ryali S, Supekar K, Abrams DA, Menon V. 2010. Sparse logistic regression for whole-brain classification of fMRI data. *Neuroimage.* 51:752–764.
- Sadagopan S, Wang X. 2010. Contribution of inhibition to stimulus selectivity in primary auditory cortex of awake primates. *J Neurosci.* 30:7314–7325.
- Schadwinkel S, Gutschalk A. 2010. Activity associated with stream segregation in human auditory cortex is similar for spatial and pitch cues. *Cereb Cortex.* 20:2863–2873.
- Schadwinkel S, Gutschalk A. 2011. Transient BOLD activity locked to perceptual reversals of auditory streaming in human auditory cortex and inferior colliculus. *J Neurophysiol.* 105:1977–1983.
- Shamma SA, Michey C. 2010. Behind the scenes of auditory perception. *Curr Opin Neurobiol.* 20:361–366.
- Shevade SK, Keerthi SS. 2003. A simple and efficient algorithm for gene selection using sparse logistic regression. *Bioinformatics.* 19:2246–2253.
- Steinschneider M, Fishman YI, Arezzo JC. 2008. Spectrotemporal analysis of evoked and induced electroencephalographic responses in primary auditory cortex (A1) of the awake monkey. *Cereb Cortex.* 18:610–625.
- Steinschneider M, Nourski KV, Kawasaki H, Oya H, Brugge JF, Howard MA III. 2011. Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. *Cereb Cortex.* 21:2332–2347.
- Striem-Amit E, Hertz U, Amedi A. 2011. Extensive cochleotopic mapping of human auditory cortical fields obtained with phase-encoding fMRI. *PLoS One.* 6:e17832.
- Sutter ML. 2000. Shapes and level tolerances of frequency tuning curves in primary auditory cortex: quantitative measures and population codes. *J Neurophysiol.* 84:1012–25.
- Sweet RA, Dorph-Petersen KA, Lewis DA. 2005. Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. *J Comp Neurol.* 491:270–289.
- Talavage TM, Sereno MI, Melcher JR, Ledden PJ, Rosen BR, Dale AM. 2004. Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. *J Neurophysiol.* 91:1282–1296.
- Tian B, Rauschecker JP. 2004. Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol.* 92:2993–3013.
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. 2001. Functional specialization in rhesus monkey auditory cortex. *Science.* 292:290–293.
- Whittingstall K, Logothetis NK. 2009. Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex. *Neuron.* 64:281–289.
- Wier CC, Jesteadt W, Green DM. 1977. Frequency discrimination as a function of frequency and sensation level. *J Acoust Soc Am.* 61:178–184.
- Woods DL, Herron TJ, Cate AD, Yund EW, Stecker GC, Rinne T, Kang X. 2010. Functional properties of human auditory cortical fields. *Front Syst Neurosci.* 3:155.
- Woods DL, Stecker GC, Rinne T, Herron TJ, Cate AD, Yund EW, Liao I, Kang X. 2009. Functional maps of human auditory cortex: effects of acoustic features and attention. *PLoS One.* 4:e5183.
- Zatorre RJ, Belin P. 2001. Spectral and temporal processing in human auditory cortex. *Cereb Cortex.* 11:946–953.
- Zatorre RJ, Belin P, Penhune VB. 2002. Structure and function of auditory cortex: music and speech. *Trends Cogn Sci.* 6:37–46.
- Zatorre RJ, Gandour JT. 2008. Neural specializations for speech and pitch: moving beyond the dichotomies. *Philos Trans R Soc Lond B Biol Sci.* 363:1087–1104.