



# HHS Public Access

Author manuscript

*Neuroimage*. Author manuscript; available in PMC 2015 November 01.

Published in final edited form as:

*Neuroimage*. 2014 November 1; 101: 598–609. doi:10.1016/j.neuroimage.2014.07.004.

## Functional organization of human auditory cortex: Investigation of response latencies through direct recordings

Kirill V. Nourski<sup>#1</sup>, Mitchell Steinschneider<sup>#2</sup>, Bob McMurray<sup>3</sup>, Christopher K. Kovach<sup>1</sup>, Hiroyuki Oya<sup>1</sup>, Hiroto Kawasaki<sup>1</sup>, and Matthew A. Howard III<sup>1</sup>

<sup>1</sup>Department of Neurosurgery, The University of Iowa, Iowa City, IA, 52242 USA

<sup>2</sup>Department of Neurology, Department of Neuroscience, Albert Einstein College of Medicine, New York, NY 10461, USA

<sup>3</sup>Department of Psychology, Department of Communication Sciences and Disorders, Department of Linguistics, The University of Iowa, Iowa City, IA, 52242 USA

# These authors contributed equally to this work.

### Abstract

The model for functional organization of human auditory cortex is in part based on findings in non-human primates, where the auditory cortex is hierarchically delineated into core, belt and parabelt fields. This model envisions that core cortex directly projects to belt, but not to parabelt, whereas belt regions are a major source of direct input for auditory parabelt. In humans, the posteromedial portion of Heschl's gyrus (HG) represents core auditory cortex, whereas the anterolateral portion of HG and the posterolateral superior temporal gyrus (PLST) are generally interpreted as belt and parabelt, respectively. In this scheme, response latencies can be hypothesized to progress in serial fashion from posteromedial to anterolateral HG to PLST. We examined this hypothesis by comparing response latencies to multiple stimuli, measured across these regions using simultaneous intracranial recordings in neurosurgical patients. Stimuli were 100 Hz click trains and the speech syllable /da/. Response latencies were determined by examining event-related band power in the high gamma frequency range. The earliest responses in auditory cortex occurred in posteromedial HG. Responses elicited from sites in anterolateral HG were neither earlier in latency from sites on PLST, nor more robust. Anterolateral HG and PLST exhibited some preference for speech syllable stimuli compared to click trains. These findings are not supportive of a strict serial model envisioning principal flow of information along HG to PLST. In contrast, data suggest that a portion of PLST may represent a relatively early stage in the auditory cortical hierarchy.

---

© 2014 Elsevier Inc. All rights reserved.

Corresponding Author: Kirill V. Nourski, MD, PhD, Department of Neurosurgery, The University of Iowa, 200 Hawkins Dr. 1815 JCP, Iowa City, IA 52242-1061 USA, Phone: +1 (319) 335-7049, Fax: +1 (319) 335-6605, kirill-nourski@uiowa.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Introduction

Sound information processing is a complex task of critical importance. The identity of sound objects and their spatial location in the environment must be inferred from a complex constellation of auditory cues (e.g., McMurray and Jongman, 2011). In order to understand this basic biological process, it is crucial to know the flow of neural activity within different auditory processing regions of the brain over time. Experimental animal models have proven extremely valuable in delineating basic organization patterns of auditory cortex (e.g. Hackett, 2007). To examine the neural processes that subserve uniquely human speech and language capabilities, it is necessary to complement this approach with the study of brain activity in human subjects, as distinct functional demands of the auditory system in humans may demand differences in brain structure.

Auditory cortex in humans occupies dorsal and lateral aspects of the superior temporal gyrus (STG). While it is clearly comprised of multiple fields, the exact number, anatomical locations and functional properties of these fields are poorly understood (e.g. Hackett, 2007). Non-human primate models suggest a framework in which the auditory cortex is hierarchically organized into core, belt and parabelt regions, subdivided into as many as thirteen areas (Rauschecker et al., 1995; Hackett et al., 1998; Brugge and Howard, 2002; Kaas and Hackett, 2005) (Fig. 1a). According to this classic model, primary auditory cortex (AI) and adjacent cortex (areas R and RT) form the core region, surrounded by belt, and then parabelt regions. Neuroanatomical tracing studies demonstrate that each area has strong bidirectional connectivity with adjacent areas such that core areas are directly interconnected with their neighboring belt, but have sparse interconnections with parabelt areas (for review, see Jones, 2003; Hackett et al., 2014). More complex connectivity patterns are currently being characterized that refine this simplified scheme, including feedforward projections from parabelt to belt (Hackett et al., 2014). Further confounding the classic serial core-belt-parabelt model of auditory cortical connectivity are the complex parallel pathways emanating from the medial geniculate complex (Jones, 2003).

Despite the greater complexity in connectivity, physiological response patterns support the hierarchical organizational model. Studies demonstrate that core areas respond robustly to pure tones and most other sounds, while higher order belt and parabelt regions of auditory cortex exhibit a progressive decrease in response to pure tones and a progressive increase and selectivity of responses to complex sounds (including conspecific vocalizations) and sound patterns (e.g., Rauschecker and Scott, 2009; Leaver and Rauschecker, 2010; Chevillet et al., 2011). It could be predicted on anatomical grounds that core areas should exhibit the shortest onset latencies, while belt and parabelt areas should be characterized by progressively longer latencies. This prediction has been confirmed in non-human primates and has been refined by demonstrating that caudal belt and parabelt areas have shorter onset latencies than equivalent rostral areas (Kajikawa et al., 2005; Camalier et al., 2012). Similar complexities reflecting both serial as well as prominent parallel processing components in the organization of core and non-core areas have been observed in anatomical and physiological studies in the cat model (e.g. Eggermont, 1998; Stecker et al., 2003; Winer, 2010).

Currently it is unclear how the core-belt-parabelt model is reflected in the organization of human auditory cortex (Fig. 1b, 1c). The superior temporal plane forms the dorsal surface of the human temporal lobe and has complex gross anatomical features (Zilles et al., 1997; Destrieux et al., 2010). Approximately half of the cortical gray matter of the superior temporal plane lines sulci extending deep into the dorsal temporal lobe, and there is considerable gross anatomical variability across individuals and between hemispheres in the same individual. These complexities, the methodological challenges associated with performing research in human subjects, and the different functional demands on the human auditory system (e.g., speech, music) make it difficult to delineate the multi-field human auditory cortex model with the same level of precision as has been achieved with experimental animals (Hackett, 2007).

The results of most anatomical, imaging and electrophysiological studies in humans lead to the conclusion that the posteromedial portion (approximately two thirds) of Heschl's gyrus (HG) is comprised of core auditory cortex (e.g., Galaburda and Sanides, 1980; Liégeois-Chauvel et al., 1991; Talavage et al., 2000; Hackett et al., 2001; Morosan et al., 2001; Woods et al., 2009). In contrast, the human homologs of non-core fields remain controversial. The anterolateral third of HG has been variously interpreted as either core (Formisano et al., 2003; Woods et al., 2010) or belt (Kaas and Hackett, 2000; Woods et al., 2009) auditory cortex. The role of auditory cortex occupying posterolateral superior temporal gyrus (PLST) within the core-belt-parabelt hierarchical model has been even more difficult to define. Some cytoarchitectonic studies characterize this region as being comprised of belt cortex (Galaburda and Sanides, 1980; Fullerton and Pandya, 2007), others interpret it as a parabelt area (Rivier and Clarke, 1997), while others place this cortex outside of the core-belt-parabelt model altogether (Sweet et al., 2005; see also Hackett, 2007 for review). Neuroimaging studies that address the functional organization of human auditory cortex often focus on the superior temporal plane and do not include PLST in the analyses (see, e.g., Baumann et al., 2013 for review). This adds to the difficulty in placing this region within the auditory processing hierarchy, and is a significant issue given evidence that a portion of PLST may represent a relatively early stage in auditory cortical processing (Brugge et al., 2003; Nourski et al., 2013, 2014).

While a wealth of postmortem anatomical studies of human auditory cortex (e.g., Galaburda and Sanides 1980; Hackett et al., 2001; Morosan et al., 2001; Sweet et al., 2005, Fullerton and Pandya, 2007, Rivier and Clarke, 1997) offer the most precise assessments of anatomy, they cannot capture the flow of auditory information through various brain regions, and understanding this flow is crucial for parceling out early auditory processing pathways. The lack of functional measurements with high resolution both spatially and temporally makes inferences regarding functional identity of different anatomical areas difficult. Measurements of the onset latency of neural responses (e.g., electro- and magnetoencephalography [EEG and MEG]) to sound stimuli may, in theory, permit clearer inferences (see, e.g., Recanzone, 2000; Kajikawa et al., 2005; Lakatos et al., 2005; Camalier et al., 2012, for animal analogues) by capturing when different parts of the system respond to sound. However, it is difficult to relate response latency to underlying functional anatomy, given the somewhat poorer spatial resolution of these techniques.

Direct recording of high gamma (70-150 Hz) cortical activity using invasive recording techniques offer a number of advantages for studying the timing of neural responses from different auditory cortical fields in humans. High gamma activity propagates in a more spatially limited fashion than lower frequency electrophysiological components, thus providing finer spatial resolution. It is difficult to measure high gamma activity using non-invasive scalp EEG or MEG methods because of this propagation property and the large distances separating some neural sources from extracranial detectors (e.g., Millman et al., 2013; but see also Sedley et al., 2012). In contrast, invasive electrocorticographic (ECoG) recordings are obtained from electrodes in immediate proximity to the sources of high gamma activity resulting in superior signal-to-noise ratio properties (Crone et al., 2001, 2006; Ray et al., 2008; Brugge et al., 2009; Edwards et al., 2009).

In the present study, direct ECoG recordings from HG and lateral STG were used to measure and compare high gamma response latencies within different regions of human auditory cortex. Extrapolating from the non-human primate model (as shown in Fig. 1b), we posited that serial processing of auditory information would be reflected by progressively longer high gamma response latencies within posteromedial HG (core), anterolateral HG (putative belt), and then PLST (putative parabelt) cortical areas. Earlier reports provide evidence that non-core auditory cortex is preferentially activated by speech compared to non-speech stimuli (e.g., Binder et al., 2000; Woods et al., 2010, 2011). Therefore, our protocol incorporated both classes of auditory stimuli in the experimental design.

There are several limitations inherent to human intracranial recording research. The number of subjects studied is usually small, the electrode arrays cover limited regions of cortex, and there is considerable inter-subject variability in both gross anatomy and electrode coverage. In order to address these issues, we developed computational tools to anatomically reconstruct and pool auditory cortex electrode location data across subjects along with statistical techniques (linear mixed effects models) that allowed us to accurately attribute observed variances to subject differences and anatomical variables.

## Methods

### Subjects

Experiments were performed in 11 neurosurgical patient volunteers (7 male, 4 female, age 22–56 years old, median age 36 years old). The subjects had medically refractory epilepsy and were undergoing chronic invasive 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 were typically weaned from their antiepileptic medications during the monitoring period at the discretion of their treating neurologist. Experimental sessions were suspended for at least three hours if a seizure occurred, and the patient had to be alert and willing to participate for the research activities to resume.

In all participants, ECoG recordings were made from only a single hemisphere. All subjects but two had left-hemisphere language dominance, as determined by intracarotid amygdala (Wada) test results; subject R149 had bilateral language dominance, and R139 had right hemisphere language dominance. In four subjects, the electrodes were implanted on the left side, while in seven others recordings were from the right hemisphere. The side of implantation is indicated by the letter prefix of the subject code (L for left, R for right). The hemisphere of recording was language-dominant in six subjects (L140, L145, L178, L258, R139, R149) and non-dominant in five other subjects (R129, R142, R180, R186, R212).

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. All subjects were native English speakers. Intracranial recordings revealed that the auditory cortical areas on the superior temporal gyrus were not epileptic foci in any of the subjects.

## Procedure

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.

## Stimuli

Experimental stimuli were trains of acoustic clicks (used previously in Brugge et al., 2009; Nourski et al., 2013) and the synthesized consonant-vowel syllable /da/ (used previously in Steinschneider et al., 1999, 2005, 2011). Clicks were digitally generated as equally-spaced rectangular pulses (0.2 ms duration) and were presented at a rate of 100 Hz (train duration 160 ms). The speech syllable /da/ was constructed on the cascade branch of a KLSYN88a speech synthesizer (Klatt and Klatt, 1990), contained 4 formants (F1 through F4), and was 175 ms in duration. Fundamental frequency began at 120 Hz and fell linearly to 80 Hz. Steady-state formant frequencies were 700, 1200, 2500, and 3600 Hz. Onset frequencies for F1, F2 and F3 were 200, 1600 and 3000 Hz. Formant transitions were 30 ms for F1 and 40 ms for F2 and F3. F4 did not contain a formant transition. A 5 ms period of frication exciting F2-F4 preceded the onset of voicing. Further details concerning stimulus parameters and response patterns elicited by these stimuli in human auditory cortex can be found in the cited articles.

The stimuli were delivered to both ears via insert earphones (ER4B, Etymotic Research, Elk Grove Village, IL) that were integrated into custom-fit earmolds. The stimuli were presented at a comfortable level (mean = 67 dB SPL, SD = 5.3 dB SPL). In each subject, the intensity difference between click trains and speech syllables was within 10 dB. Relationship between stimulus intensity and response latency was addressed in a control study (Supplementary Fig. 3). Inter-stimulus interval was chosen randomly within a Gaussian distribution (mean interval 2 s; SD = 10 ms) to reduce heterodyning in the recordings secondary to power line noise. Stimulus delivery and data acquisition were controlled by a TDT RP2.1 and RX5 or RZ2 real-time processor (Tucker-Davis Technologies, Alachua, FL).

## Recordings

Recordings were made simultaneously from HG and perisylvian cortex using multicontact depth electrodes and high density subdural grid electrodes, respectively (Howard et al., 2000, Reddy et al., 2010, Nourski and Howard, in press). Hybrid depth electrode arrays (AdTech, Racine, WI) were implanted stereotactically into HG, along its longitudinal (anterolateral to posteromedial) axis (Howard et al., 1996; Reddy et al., 2010). HG depth electrodes included 4-6 platinum macro contacts, spaced 10 mm apart, and 14-15 platinum micro contacts (diameter 40  $\mu\text{m}$ ), distributed at 2–4 mm intervals between the macro contacts. In one of the subjects (R129), two depth electrodes were implanted in the superior temporal plane, both providing HG coverage.

Multicontact subdural grid electrodes (AdTech, Racine, WI) were placed over perisylvian cortex including the STG. The recording arrays consisted of 64 or 96 platinum-iridium disc electrodes (2.3 mm exposed diameter, 5 mm center-to-center spacing) arranged in an 8 $\times$ 8 or an 8 $\times$ 12 grid and embedded in a silicon membrane. A subgaleal contact was used as a reference. Recording electrodes remained in place for approximately 2 weeks under the direction of the patients' neurologists.

## Anatomical reconstruction

Reconstruction of the anatomical locations of the implanted electrodes and their mapping onto a standardized set of coordinates across subjects was performed using software developed in-house. Contact locations of the HG depth electrodes and subdural grid electrodes were first extracted from post-implantation MR and CT scans, respectively. These were then projected onto preoperative MR scans using non-linear warping. Finally, these were then projected into the standard Montreal Neurological Institute space (MNI305) using surface-based warping. These steps were carried out according to the following procedure.

**Contact localization**—Contact localization was performed separately for the HG depth electrodes and for the surface grid electrodes. For the HG depth electrodes, contact locations were manually identified using post-implantation MR images. The locations (MR volume indices) were manually transferred onto pre-implantation MR volume. Using MR instead of CT image provided an advantage for the accurate visualization of brain structures surrounding the contacts. The locations of the micro-contacts were found by applying Catmull-Rom spline interpolant. The pre-implantation MR volumes were coregistered with the MNI space using linear affine transformation and MNI coordinates for each contact were calculated.

For the grid electrodes, locations of all grid contacts were determined from a postoperative CT scan. This was accomplished by manually identifying the location of a subset of contacts in the grid on the basis of characteristic hyper-intense radiological artifacts. Identified contacts included the 4 corner contacts and 4 to 6 interior contacts. The full 64- or 96-contact grid was fitted to these locations by thin-plate-spline (TPS) warping (Bookstein, 1989; Rohr, 2001), using a negligibly small regularization parameter. Applying TPS allowed the non-linear deformation of the grid to be closely approximated. Accuracy of fitting was

evaluated by visually comparing fitted contact locations with the contact artifacts in the CT and by verifying that inter-contact spacing fell within 0.2 mm of the expected 5 mm spacing.

After the initial grid locations were determined by CT, these were further corrected using a preexplantation MR scan (the same scan used to determine the location of the HG electrodes). Because displacement of brain parenchyma related to electrode mass-effect and post-operative swelling was often difficult to evaluate accurately on the CT scan, the results of CT-based localization were compared against a T1 MR scan obtained shortly before explantation. When significant discrepancy, greater than approximately 2 mm, was observed between CT-derived contact locations and corresponding magnetic susceptibility artifacts in the MR scan, a rigid linear transform was used to adjust grid positioning on the basis of clearly identifiable electrode-related artifacts in the MR. Most typically, the corner contacts were used as control points in this transformation.

**Pre-implantation MR to preoperative MR registration**—In order to pool the data for analysis it was necessary to co-register electrode locations to a standard reference brain. However, this could not be accomplished by the application of automated image co-registration because of tissue distortion, signal degradation and the presence of susceptibility artifacts in the pre-explantation MR. The effect of image distortion was minimized through the following steps. The pre-explantation MR was registered to a pre-operative MR—which is free of such artifacts—using TPS warping. In this step, control points were manually selected from corresponding locations in the two images. Major anatomical landmarks, including the posterior commissure, anterior commissure, genu of the corpus callosum, temporal poles, and the amygdalae were used in guiding control point selection, as were any other features that visibly corresponded between the two images, such as sulcal boundaries and venules. Between 50 and 100 control points throughout the brain were typically selected in this step. Following control point selection, TPS warping was applied to generate a non-linear transformation between the two images. The warped post-operative MR was compared against the preoperative MR to gauge the accuracy of the transformation. Any necessary adjustments were made through the addition or correction of control points in regions of substantial discrepancy.

**Surface-based coregistration of coordinates to standard MNI space**—To improve the accuracy of analyses that pooled data from multiple subjects, electrode locations were mapped onto a standard brain based on individual gyral anatomy. To that end, each subject's brain surface was constructed using Freesurfer image analysis suite (<http://surfer.nmr.mgh.harvard.edu/>). Because our present aim was to identify contact locations with respect to specific cortical structures (HG and STG) in the reference space, we applied non-linear warping with mesh vertices as control points. This step used mesh points identified as HG and STG by Freesurfer's automated cortical parcellation routine, according to Freesurfer Destrieux Atlas (Destrieux et al., 2010). For depth electrode contacts, only vertices parcellated into HG were used in computing this transform, while for temporal grid contacts overlaying STG, only vertices parcellated into STG were used. This procedure derived a 3D-volume transform from the correspondence between the subjects anatomy and the reference atlas.

## Data analysis

Recorded ECoG data were filtered (1.6–1000 Hz bandpass, 12 dB/octave rolloff), amplified (20×), and digitized at a sampling rate of 2034.5 Hz (for data recorded from clinical contacts) or 12207 Hz (for data recorded from micro contacts of the HG depth electrode). Analysis of recorded responses was done by calculating the onset latency of high gamma (70-150 Hz) event-related band power (ERBP). Data analysis was performed using custom software written in MATLAB Version 7.14.0 (MathWorks, Natick, MA, USA) and R (version 2.13.1, R Development Group). Pre-processing of ECoG data included downsampling to 1 kHz, followed by removal of power line noise by an adaptive notch filtering procedure (Nourski et al., 2013).

An initial examination of auditory evoked potential (AEP) waveforms suggested that earliest responses in auditory cortex occurred in posteromedial HG (Supplementary Fig. 1). However, as expected, AEPs recorded from different auditory cortical regions were characterized by different morphology, making latency measurements complicated and across-region latency comparisons impractical. This reinforced our motivation to focus on high gamma cortical activity.

Analysis of high gamma cortical activity was performed using wavelet transforms based on complex Morlet wavelets following the approach of Oya et al. (2002). Center frequencies ranged from 70 to 150 Hz in 5 Hz increments. ERBP was calculated for each center frequency on a trial-by-trial basis, log-transformed and normalized to mean baseline power, measured for the same center frequency within a 100 to 200 ms window prior to stimulus onset. ERBP values were then averaged across trials. The wavelet constant ratio used for time-frequency analysis was defined as  $f_0/\sigma_f = 6$ , where  $f_0$  is the center frequency of the wavelet and  $\sigma_f$  is its standard deviation in the frequency domain.

Latency of high gamma response was measured within 200 ms after stimulus onset as the time at which the lower limit of high gamma ERBP 95% confidence interval exceeded 0 dB relative to the prestimulus mean and remained positive for at least 30 ms. We recognize that using this method, amplitude and latency would not be entirely independent. Specifically, lower-amplitude responses with similar variance as larger responses might be characterized by increased onset latency. However, baseline activity has a natural variance and under most reasonable models of sound detection, in order for a sound processing network to detect a significant stimulus-related change, responses needed to exceed the variance of this baseline. Thus, our use of this criterion is motivated by the demand characteristics of sound detection in general. We therefore used the lower limit of the 95% confidence interval as a conservative criterion for variance threshold to define response onset latency.

Recording sites were included in analyses based on their anatomical location (i.e., implanted in the grey matter of the HG or overlying the lateral surface of the STG) and presence of a high gamma response to at least one of the two stimuli. Anatomical location was determined by the localization of each electrode in the pre-implantation MR for each subject individually, and not based on the common MNI coordinates. Based on these criteria, a total of 144 recording sites in HG and 289 sites on PLST from the 11 subjects were included in the analyses. Throughout the manuscript, the term “STG” is used to refer to the entire lateral

exposed surface of the gyrus, whereas “PLST” is used to refer to the acoustically responsive portion of the gyrus. For left hemisphere cases, MNI  $x$ -axis coordinates ( $x_{\text{MNI}}$ ) were multiplied by  $(-1)$  to map them onto the right-hemisphere common space.

As the MNI coordinate system axes are not aligned with the orientation of the anatomical areas of interest, it was important to rotate the coordinates along these axes to allow for a more straightforward interpretation of the different dimensions in terms of the orientation of the STG. More importantly, because of this lack of alignment, raw MNI coordinates (termed  $x_{\text{MNI}}$ ,  $y_{\text{MNI}}$ ,  $z_{\text{MNI}}$  in this manuscript) were strongly collinear – movement along STG (which is diagonal relative to standard axial and coronal planes) resulted in movement in both  $y_{\text{MNI}}$  and  $z_{\text{MNI}}$ . This collinearity made statistical analyses difficult to interpret. We thus rotated the MNI coordinates along anatomical STG axes, to eliminate this collinearity and create a more transparent statistical model. To that end, the locations of each acoustically responsive site in HG and on STG were rotated such that they could be described in terms of their location relative to the long axis of the gyrus and its bounding sulci. For HG sites, this was done in the  $x_{\text{MNI}}y_{\text{MNI}}$  plane; in STG this was done in the  $y_{\text{MNI}}z_{\text{MNI}}$  plane. To accomplish this, the coordinates in these planes were first centered by subtracting the grand mean location from each individual coordinate. Next, the best fit regression line was computed relating  $x_{\text{MNI}}$  to  $y_{\text{MNI}}$  for HG and  $y_{\text{MNI}}$  to  $z_{\text{MNI}}$  for STG. The corresponding angle of rotation,  $\theta$ , was computed from the slope of that line. Finally, each set of coordinates was rotated by  $\theta$  using standard linear algebraic techniques. The results of this was that in HG, the new  $x_{\theta}$  coordinate corresponded to the position along the long axis of HG, and the new  $y_{\theta}$  corresponded to location relative to the anterior temporal sulcus (ATS) and Heschl’s sulcus (HS). In STG, the new  $y_{\theta}$  corresponded to the posterior/anterior dimension along the gyrus, and the new  $z_{\theta}$  corresponded to the location relative to the superior temporal sulcus (STS) and Sylvian fissure (SF).

Primary analyses were conducted with linear mixed effects models, implemented in the LME4 package (version 1.04; Bates and Sarkar, 2011) of R (version 2.15.0). These analyses used latency as the dependent variable, and location ( $x_{\theta}$ ,  $y_{\theta}$ ,  $z_{\theta}$ ) along with several other factors as independent variables. This approach was adopted because subjects differed in their coverage of the auditory cortex with recording arrays. Consequently, the primary variables of interest (location in auditory cortex) were confounded with subject. Traditional general linear approaches (regression and ANOVA) may thus eliminate variation due to location when subject variance is accounted for, or ignore subject variance and incorrectly attribute this to location. In contrast, a mixed effects model can fit the effect of location for each subject, and thus properly account for this shared variance. Hemisphere was examined as a between-subject variable. It was coded in terms of both language dominance (non-dominant vs. dominant) and side (left vs. right). To compute  $p$ -values for each coefficient, we used the Satterthwaite approximation for the d.f. of the coefficients in the model, as implemented in the lmerTest package of R (version 2.0.0). For the purpose of clarity, details of the specific models are described in the Results section.

## Results

### Data from representative subjects

There was a progressive increase in onset latency of response to the 100 Hz click train and speech syllable /da/ from the posteromedial portion of HG to the anterolateral HG. Surprisingly, onset latencies on PLST were shorter, and responses tended to be larger compared to anterolateral HG. This pattern is exemplified in Figure 2, which depicts high gamma responses to these stimuli from several representative sites, from a representative subject (R212). The subject was implanted with a depth electrode and a subdural grid electrode in the right (language non-dominant) hemisphere, which allowed for simultaneous recording from HG and the lateral surface of the STG, respectively (Fig. 2a). The shortest latencies (< 25 ms) for high gamma activity occurred in posteromedial HG (Fig. 2b). From this location, response latencies increased along the long axis of the HG towards its anterolateral portion (sites A through D, indicated by arrows in Fig. 2b).

In contrast to the model that envisions serial progression of activation from posteromedial to anterolateral HG to PLST (see Fig. 1b), high gamma activity recorded from PLST was characterized by intermediate latencies, shorter than those measured in the anterolateral HG but longer than those seen in posteromedial HG (sites E, F in Fig. 2). More anterior sites (e.g., G, H in Fig. 2) did not exhibit high gamma responses to the click train stimuli large enough to meet the significance criterion (i.e., the lower limit of the 95% confidence interval did not exceed 0 dB within 200 ms after stimulus onset).

Responses to /da/ exhibited a similar systematic increase in latency along the HG (Fig. 2b, right column), but showed a more extensive activation pattern within PLST, engaging sites both anterior and posterior to the foci with the shortest-latency responses. Notably, in this subject, responses from sites in anterolateral HG were neither earlier in latency from sites on PLST, nor more robust. Also, /da/ elicited larger-amplitude responses than the click train and engaged additional sites along the lateral surface of the STG (sites G, H in Fig. 2).

An examination of all of responsive sites on HG and PLST (Fig. 3) graphically illustrates the more expansive activation by the speech syllable (rather than the click train). As previously described, there was a pronounced shift in latency along HG (Fig. 3, upper panels). The sites immediately lateral to HG represent contacts that followed the curvature on the lateral surface of STG (connected by lines in the lower panels). Remarkably, despite the greater physical distance of these PLST sites from posteromedial HG relative to anterolateral HG, latencies were consistently shorter. The lower panels of Figure 3 illustrate both the extent of activation and onset latencies of sites on PLST to both stimuli. The speech syllable activated sites beyond the area that was responsive to the click train. Similar findings were observed in recordings from the left, language-dominant hemisphere in a different subject (L178; Supplementary Fig. 2).

### Group data

The consistency of response latency profiles across all subjects obtained from HG and PLST is shown in Figure 4. Here, data are plotted from all 11 subjects in standard MNI coordinate space. Latencies increased from <25 ms to over 100 ms along the long axis of HG (dotted

outline in Fig. 4a). Latency distributions on PLST were more complex (Fig. 4b). Projection of these sites onto the horizontal plane (ovals in Fig. 4a) revealed that sites with the shortest latencies (<25 ms) were often over 20 mm away from posteromedial HG. Latencies on PLST were often markedly shorter than sites on the anterolateral HG which were in closer proximity to the posteromedial short-latency sites. Additionally, numerous sites, both in HG and on PLST, exhibited responses to the syllable /da/, but not to the click train. These sites, depicted in white in the middle panels of Figure 4, were typically found in the most anterolateral aspect of HG and the most posterior or anterior portions of PLST. The more expansive activation elicited by /da/ occurs despite the fact that both sounds were of similar duration, intensity and fundamental frequency.

A potential confound is that intensities were varied based on the subjects' comfort level that could change between recording sessions, both within and across subjects. Since this would bias the primary measure of interest (Lee et al., 1984; Howard et al., 2000), this potential tradeoff was estimated in a control study in three additional subjects (R131, R136, L175). Click trains of five rectangular pulses (100 Hz rate) were presented at multiple intensities with 5-10 dB intervals. Findings demonstrated that at typical presentation levels used in the main experiment, the latency-intensity curves had reached asymptote (Supplementary Fig. 3).

The primary question was whether there was a systematic change in latency along HG and STG. To that end, the location of each recording site was calculated on the rotated axis ( $x_0$ , for HG and  $y_0$  for STG), as described in Methods. For ease of visualization, HG and STG sites were divided into three equal-width groups according to their location along their respective axis ( $x_0$ , for HG and  $y_0$  for STG) (Fig. 5). These groups were based on location along the respective physical dimensions (not number of electrodes, e.g., tertiles). This division is not to imply any discrete cortical organization, and all of the statistical analyses presented below treat location as a continuous measure. As Figure 5a shows, median latencies to both the click train and /da/ increased along HG. On STG, latencies were longer than those in the medial two-thirds of HG, yet shorter than in the anterolateral third of HG. Additionally, latencies to the click train exhibited a slight U-shaped function with longer latencies in posterior and anterior thirds of PLST.

As noted earlier, some sites in HG and on PLST exhibited responses to the click train, but not to the speech stimulus, while others responded to the speech stimulus, but not to the click train. These differences in responses to the two stimuli across cortical regions are summarized in Figure 5b. A preference to one stimulus over the other was more common in the anterolateral third of HG and on PLST (presumptive non-core areas) compared to the medial two thirds of HG (core auditory cortex).

To examine the relationships between recording site location and response latency in more depth, separate linear mixed effects analyses were conducted on the HG and STG latencies.

### Latencies along HG

The first analysis examined the relation between latency and location along HG. A number of fixed and random factors were tested. The fixed factors included  $x_0$  (location along the

long axis of HG, centered),  $y_{\theta}$  (orthogonal axis corresponding to position relative to the ATS and HS, centered), *stimulus-class* (click train or /da/). Subject was the only random effect. However,  $x_{\theta}$  and  $y_{\theta}$  were correlated with subject due to differences in individual anatomy and electrode coverage. Therefore, this covariation required a statistical model that captured the effect of location for each subject individually, that is, random slopes of location on subject.

Before each statistical analysis, the random effects structure that best fit the data was first determined without examining any of the fixed effects. The resulting model used  $x_{\theta}$ ,  $y_{\theta}$ , *stimulus-class* and their interactions as fixed effects, and a random slope of  $x_{\theta}$  on subject.

The results of this analysis are shown in Table 1 and Figures 6a and 7a. This analysis revealed a significant main effect of  $x_{\theta}$  ( $p < 0.0001$ ), reflecting an increase in latency as location moved from posteromedial to anterolateral along HG (Fig. 6a, solid lines). *Stimulus-class* was not significant; however, there was an interaction of *stimulus-class* with  $x_{\theta}$ , such that the effect of location was stronger for the click train than /da/ ( $p < 0.0001$ ). Lastly, there was a main effect of  $y_{\theta}$  ( $p = 0.0021$ ), such that there were longer latencies as the location of the recording site moved across HG, from HS towards ATS. Figure 7a shows the predicted latencies generated by this statistical model for the area within the  $x_{\theta}y_{\theta}$  plane bound by the electrode coverage in all subjects.

The effect of hemisphere was explored with this model by adding a between-subject term reflecting whether the recorded hemisphere was language-dominant, along with all of the interactions, to the prior model. This did not result in a better fit than the previous model. We also ran this model coding hemisphere by left/right and found no differences. This suggests that the primary effects of location and *stimulus-class* were comparable in both hemispheres.

### PLST latencies

A similar approach was used to analyze the data from PLST. Here, the preliminary investigation suggested that the relation between location along the lateral surface of the gyrus ( $y_{\theta}$ ) and latency may be U-shaped rather than linear (Fig. 6b). Thus, this model included both  $y_{\theta}$  and its quadratic effect  $y_{\theta}^2$  (centered) along with the orthogonal dimension,  $z_{\theta}$  (approximate location relative to STS and SF, also centered) and *stimulus-class*. Random slopes of  $y_{\theta}$ ,  $y_{\theta}^2$  and  $z_{\theta}$  (but not their interactions) were used, as these yielded the best fitting model that also converged. Note that to achieve model convergence,  $y_{\theta}$ ,  $y_{\theta}^2$  and  $z_{\theta}$  had to be converted to Z-scores (rather than using raw centered values as in the prior analysis). Thus,  $B$  values for the location variables should be seen as the change in latency relative to 1 SD (of  $y_{\theta}$ ,  $y_{\theta}^2$  or  $z_{\theta}$ , respectively) rather than to 1 mm.

Results of this analysis are shown in Table 2, and visualizations of the predicted latencies generated by this model as a function of location are presented in Figure 7b. The main effect of  $y_{\theta}$  was not significant, indicating that there was no linear relationship with the latency along the anterior/posterior dimension of STG. In contrast, the quadratic effect ( $y_{\theta}^2$ ) was significant ( $p = 0.049$ ). This is illustrated in Figure 6b, wherein latencies tend to be the shortest in the middle third of PLST and become progressively longer at more anterior and

posterior areas. The main effect of *stimulus-class* was marginally significant ( $p = 0.062$ ), with shorter latencies for the click train compared to /da/. An interaction of  $z_{\theta}$  and  $y_{\theta}^2$  was also present ( $p = 0.0035$ ), reflecting the finding that the quadratic effect was more pronounced toward the dorsal edge of STG (i.e., closer to the SF; see Fig. 7b). Similarly, the marginal interaction of *stimulus-class* and  $y_{\theta}^2$  ( $p = 0.068$ ) was due to the fact that responses to the click train had a more pronounced curvilinear relationship with  $y_{\theta}$  (Fig. 6b, left vs. right panels). Finally, there was a significant three-way interaction of  $x_{\theta} : y_{\theta} : \textit{stimulus-class}$  ( $p = 0.0026$ ), indicating that, for /da/, the overall quadratic effect was less pronounced, and therefore less impacted by the dorsal/ventral dimension ( $z_{\theta}$ ).

As before, we ran additional models adding hemisphere (either language-dominant or left/right) and found no effects or interactions with hemisphere.

### Anterolateral HG vs Mid PLST

We previously hypothesized that a portion of PLST may represent a relatively early stage in auditory cortical processing (Nourski et al., 2013, 2014). The data presented here are consistent with this hypothesis. Mixed effects analysis demonstrated that the middle third of PLST responds faster than the surrounding tissue (see Fig. 6b, 7b). Further, previous studies also suggest that a portion of PLST may counterintuitively represent an earlier processing stage than the anterolateral HG, indicated by the morphology of the AEPs elicited by click train stimuli (Brugge et al., 2008) and the finding that PLST featured phase locking to higher-rate click trains compared to the anterolateral HG (Brugge et al., 2009; Nourski et al., 2013). However, up till now, there has been no systematic analysis of neural response latency that would provide critical support for this hypothesis.

To examine this question, a planned comparison was performed between recording sites in the lateral third of HG and the middle third of PLST. These groups of sites were divided by the same approach described in *Group data* section by dividing the range of  $x_{\theta}$  (for HG) and  $y_{\theta}$  (for STG) into three equally sized groups. Again, this is not to suggest a discrete boundary, but rather to offer a simple objective way to define the middle portion of STG. This yielded 40 and 140 sites for the lateral HG and middle PLST groups, respectively. A linear mixed effects model was employed to examine latency as a function of *brain-area* and *stimulus-class*. Random slopes of both *brain-area* and *stimulus-class* (and their interaction) were used, as these improved model fit over random intercepts alone. This model showed a highly significant effect of brain area [ $B = 16.54$ ,  $SE = 5.2$ ,  $t(8.8) = 3.2$ ,  $p=0.0116$ ] with *longer* latencies in Lat HG compared to Mid PLST (Fig. 8). The main effect of *stimulus-class* was not significant [ $B = 4.0$ ,  $SE = 7.0$ ,  $t(8.7)=0.57$ ,  $p=0.58$ ], and there was no interaction [ $B = 9.2$ ,  $SE = 9.0$ ,  $t(13.9) = 1.0$ ,  $p = 0.32$ ]. Thus, these results offer strong support to the hypothesis that the lateral third of HG does not serve as the source for early inputs into PLST.

## Discussion

This study demonstrates consistent differences in high gamma response latencies across different regions of human auditory cortex. The posteromedial two thirds of HG showed the shortest latencies, followed by the middle portion of PLST. The surrounding areas of PLST

and the anterolateral third of HG appeared to follow. This suggests a rather complex pattern of responses over time in which neural activation does not simply follow laterally down the HG to the lateral surface of PLST, but rather may feature multiple branch points outside of HG (e.g., planum temporale) or is not predominantly serial at all.

Before discussing the implications of this pattern of latencies, however, it should also be acknowledged that the physiological criteria discussed here are derived from a limited set of stimuli and are based on neural activity in patients with epilepsy. It is conceivable that this subject population might undergo long-term reorganizational changes of the temporal lobe based on their disorder. However, as discussed below, results of the current study are in accord with multiple anatomical and physiological investigations in subjects not encumbered with a neurologic disorder.

It is difficult to map the pattern of latencies observed here onto auditory field structure, as the distinction between core, belt and parabelt model does not solely derive from the sequence of activity evoked by an auditory stimulus, but also by the functional properties of these areas. Nonetheless, to the extent that the hierarchical processing represented by this model may map onto latency, the present study offers a picture that is consistent with a number of other lines of work.

The shortest latencies within the most posteromedial third of HG suggest that this region can reliably be interpreted as auditory core cortex. The middle third of HG has slightly longer latencies, and could either represent a portion of core cortex or a transition to an immediately adjacent belt region. If auditory core, the slightly longer latencies might be based on the tonotopic organization, wherein lower best frequencies evoke longer latency responses, or a core area receiving slightly later inputs from the auditory thalamus.

After the earliest responses in the posteromedial two thirds of HG, the next shortest latency responses were located within the middle portion of PLST. While latencies even in core can vary widely across areas (e.g. Camalier et al., 2012) and thus, just on latency grounds, the middle portion of PLST might be interpreted as a core area, anatomical considerations negate this idea (e.g., Hackett et al., 2001). This region may therefore represent a belt or a parabelt area. The fact that there are no electrode recording sites within the relative large expanse of superior temporal cortex located directly between core cortex on posteromedial HG and the mid-portion of PLST precludes our ability to definitively functionally characterize this intervening region, and consequently to determine the location of the core-belt, or possibly belt-parabelt boundary transitions within that region of the superior temporal plane. In support of the mid-portion of PLST being a belt area is the model of human auditory cortex as defined by fMRI (Woods et al., 2010). Compared to the mid-portion of PLST, the anterior and posterior regions of PLST have progressively longer response latencies, and the longest of all latencies were observed within anterolateral HG.

While the present study focused on measuring latency using high gamma activity, converging evidence supporting the middle portion of PLST as putative belt cortex can be found in several other measures. Functional imaging studies (e.g., Woods et al., 2010) show a region with high activation magnitudes that overlaps with the middle portion of PLST

which exhibited the shortest latencies. Strong activation in the high gamma activity elicited by tones, click trains and speech syllables also occurs in this part of PLST (Steinschneider et al., 2011; Nourski et al., 2013, 2014).

Further compelling evidence is provided by studies utilizing direct electrical stimulation of the brain. Stimulation of core auditory cortex in posteromedial HG was shown to elicit responses on PLST with latencies as short as 3 ms (Howard et al., 2000; Brugge et al., 2003). Such short latencies cannot be readily explained by activation occurring through either a corticothalamic loop or an intervening belt region. Either scenario would entail a circuit containing an additional synaptic relay that would likely delay the response by more than 3 ms due to axonal conduction and synaptic delay. Thus, these studies support a more direct connection between posteromedial HG and PLST.

Additionally, connectivity analysis of AEP data obtained simultaneously from posteromedial HG and PLST using Granger causality techniques has suggested direct functional connectivity between the two regions (Oya et al., 2007); this suggests a belt, rather than parabelt (or beyond) characterization of PLST. Finally, source localization of middle latency auditory cortical AEP and AEF components using simultaneous EEG/MEG latency measures indicates a progression of activity beginning in posteromedial HG, followed by activation of PLST and afterward, activation of anterolateral HG (Yvert et al., 2001). This sequence of activation is not consistent with models of human auditory cortex that posit that the activation of the anterolateral third of HG begins prior to PLST (Inui et al., 2005).

Supportive data are also provided by studies investigating phase locking to repetitive acoustic transients. These demonstrate progressively diminished phase locking that closely mirrors our latency findings (Eggermont, 2001; Joris et al., 2004; Wang et al., 2008). Posteromedial HG (core) exhibited phase-locked responses to click trains at rates of up to 100-200 Hz, followed by phase locking on PLST that consistently reached 50 Hz (Brugge et al., 2009; Nourski et al., 2013). In contrast, simultaneous recordings in anterolateral HG were characterized by minimal phase locking.

Finally, additional support for the middle portion of PLST as a belt region comes from an earlier ECoG study that showed that pure tones strongly activated this segment of PLST in a spatially distributed manner (Nourski et al., 2014). Suprathreshold pure tone stimuli activate belt areas of auditory cortex in the unanesthetized macaque (Petkov et al., 2006; Tanji et al., 2010). In contrast, spatially distributed responses to pure tones would not be expected to occur in a parabelt field (Chevillet et al., 2011). In this latter fMRI study, weak activation by pure tones was used as a defining criterion for a parabelt field. Overall, we conclude that response properties of PLST are not entirely consistent with those of parabelt fields and are more characteristic of that seen in belt.

While much evidence supports the proposed organizational scheme wherein a portion of PLST represents belt auditory cortex, there are several issues that remain outstanding. First and foremost, multiple core and non-core areas of auditory cortex receive thalamic inputs from different subdivisions of the medial geniculate complex (Jones and Burton, 1976;

Jones, 2003; Winer, 2010). Therefore, the earliest activity seen on PLST could potentially reflect direct activation by the auditory thalamus as opposed to intracortical connections from adjacent core or belt areas. The present study cannot unambiguously resolve this possibility. However, PLST responsivity is strongly affected by general anesthesia, whereas early activity on posteromedial HG is not (Howard et al., 2000; Nourski et al., 2009a, 2009b). If both areas are receiving parallel inputs from the auditory thalamus, it might be predicted that the earliest activity on PLST would reflect a thalamocortical projection and thus would be expected to be preserved under general anesthesia in a manner similar to that occurring on HG (Boly et al., 2012; Monti et al., 2013). It should be noted, however, that activity on the anterolateral HG appears to be more resistant to general anesthesia compared to PLST (Nourski et al., 2009a, 2009b). The reason for this discrepancy remains unclear.

Selectivity to the click train or acoustically more complex speech syllable /da/ differed between the posteromedial two thirds of HG and all other brain areas sampled. This is in concordance with findings that the human auditory core is robustly activated by either relatively simple (pure tones, click trains) or complex (speech, music) sounds, whereas non-core areas exhibit a greater selectivity for more spectrally complex sounds (e.g., Belin et al., 2000; Rauschecker and Scott, 2009; Chevillet et al., 2011). We expect that given the limited set of speech and non-speech stimuli used in the current analysis, parcellation of auditory cortex based on this criterion will be subject to later refinement.

The current physiological delineations of auditory cortex are not completely congruent with cytoarchitectonic studies. Most anatomical studies agree that the core auditory cortex occupies the posteromedial two thirds of HG (e.g., Galaburda and Sanides, 1980; Hackett et al., 2001; Morosan et al., 2001). This delineation is in accord with current findings as well as previous functional imaging and electrophysiology studies (Liegeois-Chauvel et al., 1991; Talavage et al., 2000; Woods et al., 2009). However, anterolateral HG, despite its anatomical proximity and cytoarchitectonic similarity to the auditory core (summarized in Hackett, 2007), has onset latencies longer than those on the anatomically more distant PLST. This finding is difficult to reconcile with anatomy. It is possible that our choice of stimuli failed to adequately activate anterolateral HG and that stimuli with different acoustic characteristics would have led to earlier and more robust responses. Future work will be required to address this disparity.

Anatomically, PLST has been variously described as belt, parabelt, or auditory-related cortex. Response latencies were the shortest in the middle third of PLST and surrounded by regions anterior and posterior with similar latencies that were longer than seen in the intervening area. These differences suggest the possibility that there are functional transition zones within PLST. It remains to be determined whether these transitions correspond to boundaries between belt and parabelt, or parabelt and auditory-related cortex.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

We thank Haiming Chen, Rachel Gold, Richard Reale and Ariane Rhone for help with data collection and analysis. This study was supported by NIH R01-DC04290, R01-DC00657, R01-DC008089, UL1RR024979, Hearing Health Foundation and the Hoover Fund.

## References

- Bates, D.; Sarkar, D. LME4: Linear mixed-effects models using S4 classes. 2011. (R Package Version 0.9975-11) [Computer software manual]
- Baumann S, Petkov CI, Griffiths TD. A unified framework for the organization of the primate auditory cortex. *Front Syst Neurosci.* 2013; 7:11. [PubMed: 23641203]
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. *Nature.* 2000; 403:309–312. [PubMed: 10659849]
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET. Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex.* 2000; 10:512–528. [PubMed: 10847601]
- Boly M, Moran R, Murphy M, Boveroux P, Bruno MA, Noirhomme Q, Ledoux D, Bonhomme V, Brichant JF, Tononi G, Laureys S, Friston K. Connectivity changes underlying spectral EEG changes during propofol-induced loss of consciousness. *J Neurosci.* 2012; 32:7082–7090. [PubMed: 22593076]
- Bookstein FL. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on pattern analysis and machine intelligence.* 1989; 11:567–585.
- Brugge, JF.; Howard, MA. Hearing. In: Ramachandran, VS., editor. *Encyclopedia of the Human Brain.* Academic Press; New York: 2002. p. 429-448.
- Brugge JF, Volkov IO, Garell PC, Reale RA, Howard MA III. Functional connections between auditory cortex on Heschl's gyrus and on the lateral superior temporal gyrus in humans. *J Neurophysiol.* 2003; 90:3750–3763. [PubMed: 12968011]
- Brugge JF, Volkov IO, Oya H, Kawasaki H, Reale RA, Fenoy A, Steinschneider M, Howard MA III. Functional localization of auditory cortical fields of human: click-train stimulation. *Hear Res.* 2008; 238:12–24. [PubMed: 18207680]
- Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA. Coding of repetitive transients by auditory cortex on Heschl's gyrus. *J Neurophysiol.* 2009; 102:2358–2374. [PubMed: 19675285]
- Camalier CR, D'Angelo WR, Sterbing-D'Angelo SJ, de la Mothe LA, Hackett TA. Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. *Proc Natl Acad Sci U S A.* 2012; 109:18168–18173. [PubMed: 23074251]
- Chevillet M, Riesenhuber M, Rauschecker JP. Functional correlates of the anterolateral processing hierarchy in human auditory cortex. *J Neurosci.* 2011; 31:9345–9352. [PubMed: 21697384]
- Crone NE, Boatman D, Gordon B, Hao L. Induced electrocorticographic gamma activity during auditory perception. *Clin Neurophysiol.* 2001; 112:565–582. [PubMed: 11275528]
- Crone NE, Sinai A, Korzeniewska A. High-frequency gamma oscillations and human brain mapping with electrocorticography. *Prog Brain Res.* 2006; 159:275–295. [PubMed: 17071238]
- Destrieux C, Fischl B, Dale A, Halgren E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage.* 2010; 53:1–15. [PubMed: 20547229]
- Edwards E, Soltani M, Kim W, Dalal SS, Nagarajan SS, Berger MS, Knight RT. Comparison of time-frequency responses and the event-related potential to auditory speech stimuli in human cortex. *J Neurophysiol.* 2009; 102:377–386. [PubMed: 19439673]
- Eggermont JJ. Representation of spectral and temporal sound features in three cortical fields of the cat. Similarities outweigh differences. *J Neurophysiol.* 1998; 80:2743–2764.
- Eggermont JJ. Between sound and perception: reviewing the search for a neural code. *Hear Res.* 2001; 157:1–42. [PubMed: 11470183]
- Galaburda AM, Sanides F. Cytoarchitectonic organization of the human auditory cortex. *J Comp Neurol.* 1980; 190:597–610. [PubMed: 6771305]

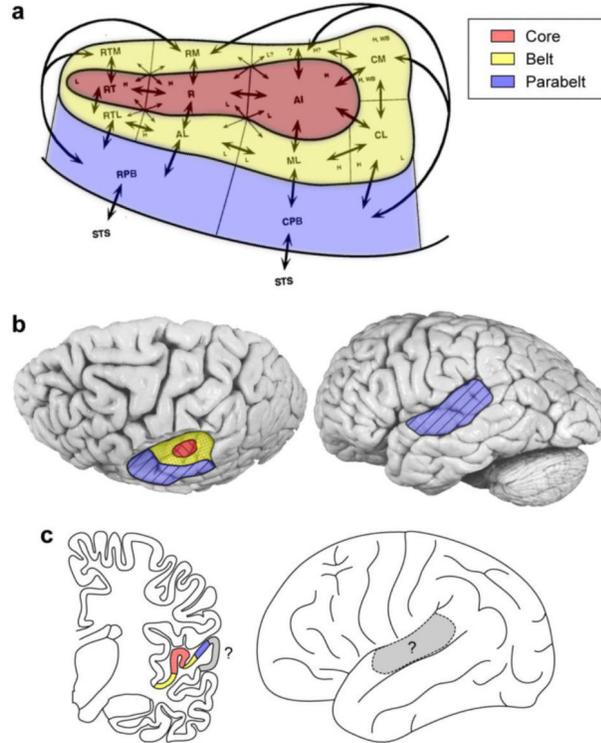
- Formisano E, Kim DS, Di Salle F, van de Moortele PF, Ugurbil K, Goebel R. Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron*. 2003; 40:859–969. [PubMed: 14622588]
- Fullerton BC, Pandya DN. Architectonic analysis of the auditory-related areas of the superior temporal region in human brain. *J Comp Neurol*. 2007; 504:470–498. [PubMed: 17701981]
- Hackett, TA. Organization and correspondence of the auditory cortex of humans and nonhuman primates. In: Kaas, JH., editor. *Evolution of Nervous Systems Vol. 4: Primates*. Academic Press; New York: 2007. p. 109-119.
- Hackett TA, Stepniewska I, Kaas JH. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *J Comp Neurol*. 1998; 394:475–495. [PubMed: 9590556]
- Hackett TA, Preuss TM, Kaas JH. Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J Comp Neurol*. 2001; 441:197–222. [PubMed: 11745645]
- Hackett TA, de la Mothe LA, Camalier CR, Falchier A, Lakatos P, Kajikawa Y, Schroeder CE. Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: refining the hierarchical model. *Front Neurosci*. 2014; 8:72. [PubMed: 24795550]
- Howard MA, Volkov IO, Abbas PJ, Damasio H, Ollendieck MC, Granner MA. A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. *Brain Res*. 1996; 724:260–264. [PubMed: 8828578]
- Howard MA, Volkov IO, Mirsky R, Garell PC, Noh MD, Granner M, Damasio H, Steinschneider M, Reale RA, Hind JE, Brugge JF. Auditory cortex on the posterior superior temporal gyrus of human cerebral cortex. *J Comp Neurol*. 2000; 416:76–92.
- Inui K, Okamoto H, Miki K, Gunji A, Kakigi R. Serial and parallel processing in the human auditory cortex: a magnetoencephalographic study. *Cereb Cortex*. 2006; 16:18–30. [PubMed: 15800024]
- Jones EG. Chemically defined parallel pathways in the monkey auditory system. *Ann N Y Acad Sci*. 2003; 999:218–33. [PubMed: 14681146]
- Jones EG, Burton H. Areal differences in the laminar distribution of thalamic afferents in cortical fields of the insular, parietal and temporal regions of primates. *J Comp Neurol*. 1976; 168:197–247. [PubMed: 821974]
- Joris PX, Schreiner CE, Rees A. Neural processing of amplitude-modulated sounds. *Physiol Rev*. 1994; 84:541–577. [PubMed: 15044682]
- Kaas JH, Hackett TA. Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci U S A*. 2000; 97:11793–11799. [PubMed: 11050211]
- Kaas, JH.; Hackett, TA. Subdivisions and connections of auditory cortex in primates: A working model. In: Konig, R.; Heil, P.; Budinger, E.; Scheich, H., editors. *Auditory cortex. A Synthesis of Human and Animal Research*. Vol. 2005. Lawrence Erlbaum Associates; Mahwah, NJ: 2005. p. 7-26.
- Kajikawa Y, de La Mothe L, Blumell S, Hackett TA. A comparison of neuron response properties in areas A1 and CM of the marmoset monkey auditory cortex: tones and broadband noise. *J Neurophysiol*. 2005; 93:22–34. [PubMed: 15342713]
- Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am*. 1990; 87:820–857. [PubMed: 2137837]
- Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J Neurophysiol*. 2005; 94:1904–1911. [PubMed: 15901760]
- Leaver AM, Rauschecker JP. Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. *J Neurosci*. 2010; 30:7604–7612. [PubMed: 20519535]
- Lee YS, Lueders H, Dinner DS, Lesser RP, Hahn J, Klem G. Recording of auditory evoked potentials in man using chronic subdural electrodes. *Brain*. 1984; 107:115–131. [PubMed: 6697149]
- Liégeois-Chauvel C, Musolino A, Chauvel P. Localization of the primary auditory area in man. *Brain*. 1991; 114:139–151. [PubMed: 1900211]
- McMurray B, Jongman A. What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations. *Psychol Rev*. 2011; 118:219–246. [PubMed: 21417542]

- Millman RE, Prendergast G, Hymers M, Green GG. Representations of the temporal envelope of sounds in human auditory cortex: can the results from invasive intracortical “depth” electrode recordings be replicated using non-invasive MEG “virtual electrodes”? *Neuroimage*. 2013; 64:185–196. [PubMed: 22989625]
- Monti MM, Lutkenhoff ES, Rubinov M, Boveroux P, Vanhaudenhuyse A, Gosseries O, Bruno MA, Noirhomme Q, Boly M, Laureys S. Dynamic change of global and local information processing in propofol-induced loss and recovery of consciousness. *PLoS Comput Biol*. 2013; 9:e1003271. [PubMed: 24146606]
- Morosan P, Rademacher J, Schleicher A, Amunts K, Schormann T, Zilles K. Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage*. 2001; 13:684–701. [PubMed: 11305897]
- Nourski KV, Howard MA III. Invasive recordings in human auditory cortex. *Hand Clin Neurol*. (in press).
- Nourski, K.; Reddy, C.; Oya, H.; Kawasaki, H.; Brugge, J.; Howard, M. Differential effects of general anesthesia on auditory cortical fields in humans. 15th Annual Meeting of the Organization for Human Brain Mapping; June 18 – 22, 2009; San Francisco, CA. 2009a.
- Nourski, K.; Reddy, C.; Oya, H.; Kawasaki, H.; Brugge, J.; Howard, M. General anesthesia differentially affects responses from auditory cortical fields in humans. 3rd International Conference on Auditory Cortex; Magdeburg, Germany. August 29 – September 2, 2009; 2009b. p. 116
- Nourski KV, Brugge JF, Reale RA, Kovach CK, Oya H, Kawasaki H, Jenison RL, Howard MA III. Coding of repetitive transients by auditory cortex on posterolateral superior temporal gyrus in humans: an intracranial electrophysiology study. *J Neurophysiol*. 2013; 109:1283–1295. [PubMed: 23236002]
- Nourski KV, Steinschneider M, Oya H, Kawasaki H, Jones RD, Howard MA III. Spectral organization of the human lateral superior temporal gyrus revealed by intracranial recordings. *Cereb Cortex*. 2014; 24:340–352. [PubMed: 23048019]
- Oya H, Kawasaki H, Howard MA III, Adolphs R. Electrophysiological responses in the human amygdala discriminate emotion categories of complex visual stimuli. *J Neurosci*. 2002; 22:9502–9512. [PubMed: 12417674]
- Oya H, Poon PW, Brugge JF, Reale RA, Kawasaki H, Volkov IO, Howard MA III. Functional connections between auditory cortical fields in humans revealed by Granger causality analysis of intra-cranial evoked potentials to sounds: comparison of two methods. *Biosystems*. 2007; 89:198–207. [PubMed: 17184906]
- Petkov CI, Kayser C, Augath M, Logothetis NK. Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biol*. 2006; 4:e215. [PubMed: 16774452]
- Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci*. 2009; 12:718–724.
- Rauschecker JP, Tian B, Hauser M. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*. 1995; 268:111–114. [PubMed: 7701330]
- Ray S, Niebur E, Hsiao SS, Sinai A, Crone NE. High-frequency gamma activity (80-150Hz) is increased in human cortex during selective attention. *Clin Neurophysiol*. 2008; 119:116–133. [PubMed: 18037343]
- Recanzone GH. Response profiles of auditory cortical neurons to tones and noise in behaving macaque monkeys. *Hear Res*. 2000; 150:104–118. [PubMed: 11077196]
- Reddy CG, Dahdaleh NS, Albert G, Chen F, Hansen D, Nourski K, Kawasaki H, Oya H, Howard MA III. A method for placing Heschl gyrus depth electrodes. *J Neurosurg*. 2010; 112:1301–1307. [PubMed: 19663547]
- Rivier F, Clarke S. Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. *Neuroimage*. 1997; 6:288–304. [PubMed: 9417972]
- Rohr K, Stiehl HS, Sprengel R, Buzug TM, Weese J, Kuhn MH. Landmark-based elastic registration using approximating thin-plate splines. *IEEE Transactions on medical imaging*. 2001; 20:526–534. [PubMed: 11437112]

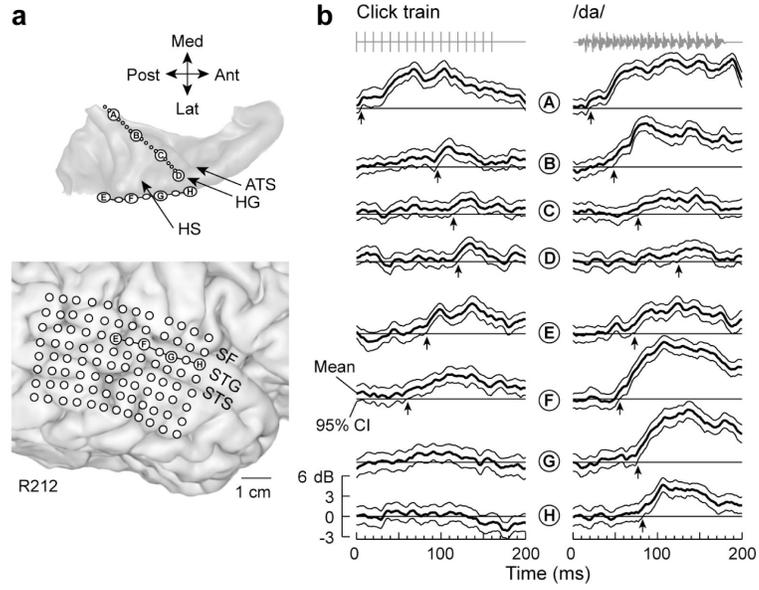
- Sedley W, Teki S, Kumar S, Overath T, Barnes GR, Griffiths TD. Gamma band pitch responses in human auditory cortex measured with magnetoencephalography. *Neuroimage*. 2012; 59:1904–1911. [PubMed: 21925281]
- Stecker GC, Mickey BJ, Macpherson EA, Middlebrooks JC. Spatial sensitivity in field PAF of cat auditory cortex. *J Neurophysiol*. 2003; 89:2889–2903. [PubMed: 12611946]
- Steinschneider M, Volkov IO, Noh MD, Garell PC, Howard MA III. Temporal encoding of the voice onset time phonetic parameter by field potentials recorded directly from human auditory cortex. *J Neurophysiol*. 1999; 82:2346–57. [PubMed: 10561410]
- Steinschneider M, Volkov IO, Fishman YI, Oya H, Arezzo JC, Howard MA III. Intracortical responses in human and monkey primary auditory cortex support a temporal processing mechanism for encoding of the voice onset time phonetic parameter. *Cereb Cortex*. 2005; 15:170–186. [PubMed: 15238437]
- Steinschneider M, Nourski KV, Kawasaki H, Oya H, Brugge JF, Howard MA III. Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. *Cereb Cortex*. 2011; 21:2332–2347. [PubMed: 21368087]
- Sweet RA, Dorph-Petersen KA, Lewis DA. Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. *J Comp Neurol*. 2005; 491:270–289. [PubMed: 16134138]
- Talavage TM, Ledden PJ, Benson RR, Rosen BR, Melcher JR. Frequency-dependent responses exhibited by multiple regions in human auditory cortex. *Hear Res*. 2000; 150:225–244. [PubMed: 11077206]
- Tanji K, Leopold DA, Ye FQ, Zhu C, Malloy M, Saunders RC, Mishkin M. Effect of sound intensity on tonotopic fMRI maps in the unanesthetized monkey. *Neuroimage*. 2010; 49:150–7. [PubMed: 19631273]
- Wang X, Lu T, Bendor D, Bartlett E. Neural coding of temporal information in auditory thalamus and cortex. *Neuroscience*. 2008; 157:484–494. [PubMed: 19143093]
- Winer, JA. A profile of auditory forebrain connections and circuits. In: Winer, JA.; Schreiner, CE., editors. *The Auditory Cortex*. Springer; New York: 2010. p. 41-74.
- Woods DL, Stecker GC, Rinne T, Herron TJ, Cate AD, Yund EW, Liao I, Kang X. Functional maps of human auditory cortex: effects of acoustic features and attention. *PLoS One*. 2009; 4:e5183. [PubMed: 19365552]
- Woods DL, Herron TJ, Cate AD, Yund EW, Stecker GC, Rinne T, Kang X. Functional properties of human auditory cortical fields. *Front Syst Neurosci*. 2010; 4:155. [PubMed: 21160558]
- Woods DL, Herron TJ, Cate AD, Kang X, Yund EW. Phonological processing in human auditory cortical fields. *Front Hum Neurosci*. 2011; 5:42. [PubMed: 21541252]
- Yvert B, Crouzeix A, Bertrand O, Seither-Preisler A, Pantev C. Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. *Cereb Cortex*. 2001; 11:411–423. [PubMed: 11313293]
- Zilles K, Schleicher A, Langemann C, Amunts K, Morosan P, Palomero-Gallagher N, Schormann T, Mohlberg H, Bürgel U, Steinmetz H, Schlaug G, Roland PE. Quantitative analysis of sulci in the human cerebral cortex: development, regional heterogeneity, gender difference, asymmetry, intersubject variability and cortical architecture. *Hum Brain Mapp*. 2007; 5:218–221. [PubMed: 20408218]

### Highlights

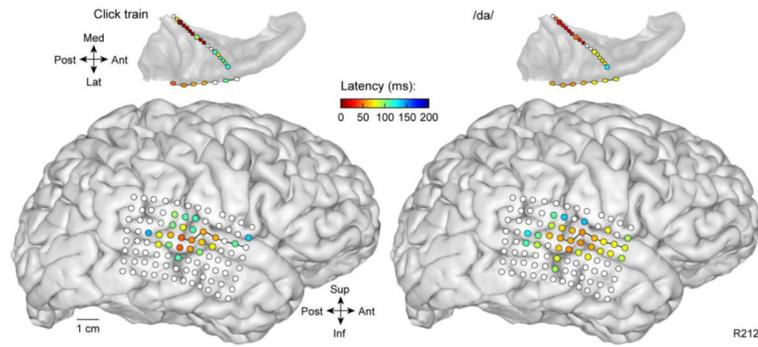
- We measured latencies of responses to sounds in multiple auditory cortical regions
- Earliest responses were found in posteromedial Heschl's gyrus (HG) (auditory core)
- Anterolateral HG was characterized by the longest latencies
- Posterolateral superior temporal gyrus (PLST) had intermediate latencies
- Part of PLST may represent a relatively early stage in auditory cortical hierarchy



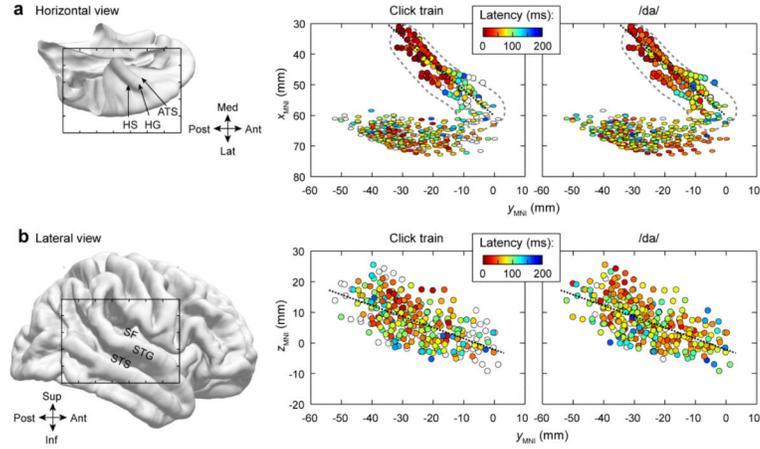
**Fig. 1.** Core-belt-parabelt hierarchical model of primate auditory cortex. **a:** A summary of primate auditory cortex subdivisions and connections. Modified from Hackett et al. (1998). Arrows indicate directions and prominence of connections. Tonotopic gradients within areas are indicated by H (high frequency), L (low frequency) and WB (wideband). **b, c:** Two contemporary models of human auditory cortex parcellation (modified from Brugge and Howard, 2002, and Hackett, 2007, respectively). Core, belt and parabelt areas are shown in red, yellow and blue, respectively.



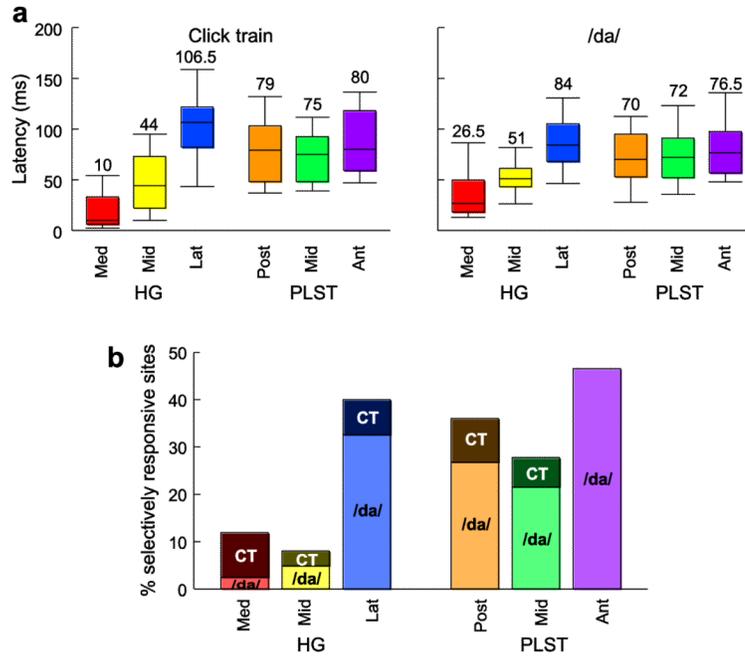
**Fig. 2.** High gamma responses recorded from the human auditory cortex. **a:** Location of recording contacts in the superior temporal plane (top) and perisylvian cortex (bottom) in a representative subject (R212). **b:** High gamma ERBP recorded from eight representative sites (A through H, see panel **a**) in response to the 100 Hz click train (left) and speech syllable /da/ (right). Thick and thin lines correspond to cross-trial mean ERBP and its 95% confidence interval (CI), respectively. Arrows indicate measured high gamma response latency.



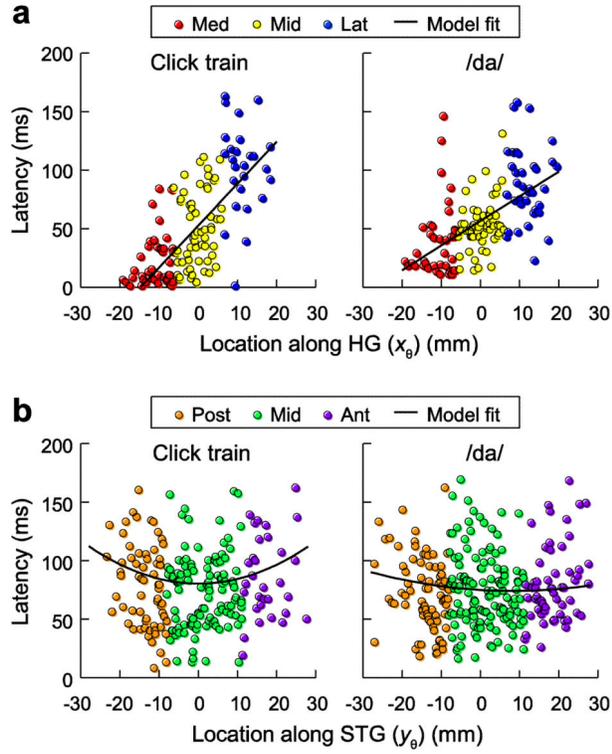
**Fig. 3.** Summary of high gamma ERBP latency measurements from the right (non-dominant) hemisphere in a representative subject (R212). Superior temporal plane and lateral view of the right hemisphere are shown in top and bottom, respectively. Response latencies to the 100 Hz click train and speech syllable /da/ are summarized in left and right panels, respectively.



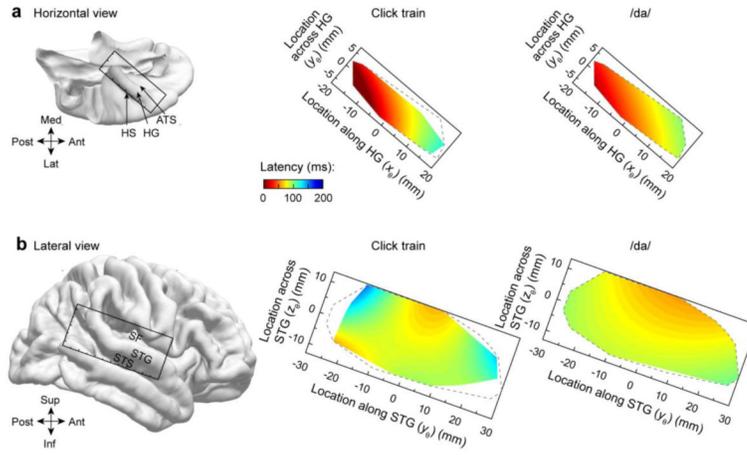
**Fig. 4.** Summary of latency measurements across all 11 subjects. Locations of responsive sites are plotted in MNI coordinate space. FreeSurfer average template brain is shown in the left panels. Color-coded latency values of high gamma responses to the 100 Hz click train and speech syllable /da/ are presented in middle and right panels, respectively. **a:** Horizontal view of HG depth electrode contacts (circles) and STG subdural grid contacts (ovals). Grey dashed line represents the location of HG in superior temporal plane. **b:** Lateral view of STG subdural grid contacts (same sites as depicted by ovals in **a**). Dotted lines represent axes of coordinate rotation for HG and PLST sites ( $x_{\theta}$  and  $y_{\theta}$ , respectively).



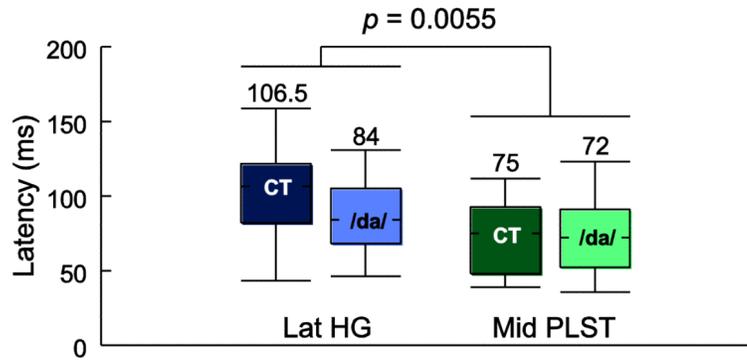
**Fig. 5.** Comparison of response properties across different regions of HG and PLST. **a:** Responses to the click train and /da/ are characterized in left and right panels, respectively. Box plots show medians, quartiles, 5th and 95th percentiles. Values above error bars indicate median latencies in ms. Note that the division of regions is based on gross anatomical criteria, but is for visualization only and is not intended to reflect any specific field boundaries. **b:** Proportions of sites that selectively responded to either the click train (CT) but not /da/ or vice versa are plotted for different regions within HG and PLST (100% corresponds to all responsive sites; the remainder represents sites responding to both stimuli).



**Fig. 6.** Statistical analysis of high gamma latency distributions in human auditory cortex. Responses to the click train and /da/ are characterized in the panels on the left and right sides, respectively. **a:** Response latencies in HG. Data from medial, middle and lateral portions of HG are shown in red, yellow and blue, respectively. **b:** Response latencies on PLST. Data from posterior, middle and anterior portions of PLST are shown in orange, green and purple, respectively. Solid lines show statistical model predictions of response latencies.



**Fig. 7.** Statistical model predictions of response latencies plotted in centered and rotated coordinate axes. **a:** Predicted response latencies in HG. **b:** Predicted response latencies on PLST. FreeSurfer template brain is shown on the left. Predicted response latencies to the click train and /da/ are shown in middle and right columns, respectively. Model predictions are bound by the convex envelopes of locations of responsive sites.



**Fig 8.** Comparison of response latencies to the click train (CT) and /da/ measured in lateral portion of HG (Lat HG) and middle portion of PLST (Mid PLST).

**Table 1**

Results of a linear mixed model for HG recordings. The model examined latency as a function of location in the  $x_{\theta}$   $y_{\theta}$  plane and *stimulus-class*.

|   | <i>B</i> | <i>SE</i> | <i>t</i> | <i>df</i> | <i>p</i>     |
|---|----------|-----------|----------|-----------|--------------|
| Intercept   | 54.68    | 5.76      | 9.5      | 8.6       | <0.0001 **** |
| $x_{\theta}$  | 2.84     | 0.39      | 7.3      | 8.7       | <0.0001 **** |
| $y_{\theta}$  | -4.16    | 1.33      | -3.1     | 148.1     | 0.0021 **    |
| <i>stimulus-class</i>                               | -3.93    | 2.92      | -1.3     | 234.3     | 0.18         |
| $x_{\theta} : y_{\theta}$                           | 0.17     | 0.11      | 1.5      | 95.1      | 0.13         |
| $x_{\theta} : \textit{stimulus-class}$              | 1.47     | 0.34      | 4.3      | 235.9     | <0.0001 **** |
| $y_{\theta} : \textit{stimulus-class}$              | -1.16    | 1.38      | -0.8     | 233.8     | 0.40         |
| $x_{\theta} : y_{\theta} : \textit{stimulus-class}$ | 0.20     | 0.15      | 1.3      | 235.6     | 0.20         |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

Results of a linear mixed model for STG recordings. Model examined latency as a function of location in the  $y_{\theta}$   $z_{\theta}$  plane and *stimulus-class* ( $p$ -values greater than 0.2 are not shown).

|   | <i>B</i> | <i>SE</i> | <i>t</i> | <i>df</i> | <i>p</i>      |
|---|----------|-----------|----------|-----------|---------------|
| Intercept   | 77.55    | 5.45      | 14.2     | 7.3       | 0.000001 **** |
| $y_{\theta}$  | 0.11     | 4.32      | 0.0      | 7.6       |               |
| $y_{\theta}^2$  | 7.01     | 2.86      | 2.5      | 6         | 0.0497 *      |
| $z_{\theta}$  | -6.31    | 4.05      | -1.6     | 8.5       | 0.156         |
| <i>stimulus-class</i>                                 | 4.58     | 2.45      | 1.9      | 441.4     | 0.062         |
| $y_{\theta} : z_{\theta}$                             | -3.42    | 1.94      | -1.8     | 164.9     | 0.080         |
| $y_{\theta}^2 : z_{\theta}$                           | 5.47     | 1.86      | 2.9      | 331.8     | 0.0035 **     |
| $y_{\theta} : \textit{stimulus-class}$                | 2.02     | 2.65      | 0.8      | 438.1     |               |
| $y_{\theta}^2 : \textit{stimulus-class}$              | 5.15     | 2.81      | 1.8      | 438.4     | 0.068         |
| $z_{\theta} : \textit{stimulus-class}$                | 0.28     | 2.64      | 0.1      | 441.2     |               |
| $y_{\theta} : z_{\theta} : \textit{stimulus-class}$   | -3.03    | 2.94      | -1.0     | 439.7     |               |
| $y_{\theta}^2 : z_{\theta} : \textit{stimulus-class}$ | 10.06    | 3.32      | 3.0      | 442.1     | 0.0026 **     |