Behavioral/Cognitive

Gamma Activation and Alpha Suppression within Human Auditory Cortex during a Speech Classification Task

Kirill V. Nourski,[1,2*] Mitchell Steinschneider,[3*] Ariane E. Rhone,[1] Christopher K. Kovach,[1] Hiroto Kawasaki,[1] and Matthew A. Howard III[1,2,4]

[1]Department of Neurosurgery, University of Iowa, Iowa City, Iowa 52242, [2]Iowa Neuroscience Institute, University of Iowa, Iowa City, Iowa 52242, [3]Departments of Neurology and Neuroscience, Albert Einstein College of Medicine, Bronx, New York 10461, and [4]Pappajohn Biomedical Institute, University of Iowa, Iowa City, Iowa 52242

The dynamics of information flow within the auditory cortical hierarchy associated with speech processing and the emergence of hemispheric specialization remain incompletely understood. To study these questions with high spatiotemporal resolution, intracranial recordings in 29 human neurosurgical patients of both sexes were obtained while subjects performed a semantic classification task. Neural activity was recorded from posteromedial portion of Heschl's gyrus (HGPM) and anterolateral portion of Heschl's gyrus (HGAL), planum temporale (PT), planum polare, insula, and superior temporal gyrus (STG). Responses to monosyllabic words exhibited early gamma power increases and a later suppression of alpha power, envisioned to represent feedforward activity and decreased feedback signaling, respectively. Gamma activation and alpha suppression had distinct magnitude and latency profiles. HGPM and PT had the strongest gamma responses with shortest onset latencies, indicating that they are the earliest auditory cortical processing stages. The origin of attenuated top-down influences in auditory cortex, as indexed by alpha suppression, was in STG and HGAL. Gamma responses and alpha suppression were typically larger to nontarget words than tones. Alpha suppression was uniformly greater to target versus nontarget stimuli. Hemispheric bias for words versus tones and for target versus nontarget words, when present, was left lateralized. Better task performance was associated with increased gamma activity in the left PT and greater alpha suppression in HGPM and HGAL bilaterally. The prominence of alpha suppression during semantic classification and its accessibility for noninvasive electrophysiologic studies suggests that this measure is a promising index of auditory cortical speech processing.

Key words: **electrocorticography; Heschl's gyrus; iEEG; insula; superior temporal gyrus; superior temporal plane**

Significance Statement

Understanding the dynamics of cortical speech processing requires the use of active tasks. This is the first comprehensive intracranial electroencephalography study to examine cortical activity within the superior temporal plane, lateral superior temporal gyrus, and the insula during a semantic classification task. Distinct gamma activation and alpha suppression profiles clarify the functional organization of feedforward and feedback processing within the auditory cortical hierarchy. Asymmetries in cortical speech processing emerge at early processing stages. Relationships between cortical activity and task performance are interpreted in the context of current models of speech processing. Results lay the groundwork for iEEG studies using connectivity measures of the bidirectional information flow within the auditory processing hierarchy.

Introduction

The dynamics of speech processing within auditory cortex and the emergence of hemispheric asymmetries remain incompletely understood (Hickok and Poeppel, 2015; Chang et al., 2015). Auditory cortex in posteromedial portion of Heschl's gyrus (HGPM) and anterolateral portion of Heschl's gyrus (HGAL), posterior portion of the lateral superior temporal gyrus (STGP), and middle portion of the lateral superior temporal gyrus (STGM) have been extensively studied using intracranial electroencephalography (iEEG; Crone et al., 2001; Bitterman et al., 2008; Mesgarani and Chang, 2012; Nourski et al., 2014; Patel et al., 2018). Only recently has iEEG been used to study physiologic

Received Nov. 2, 2021; revised Jan. 11, 2022; accepted Apr. 22, 2022.

Author contributions: K.V.N., M.S., A.E.R., and M.A.H. designed research; K.V.N., A.E.R., and H.K. performed research; K.V.N. and C.K.K. contributed unpublished reagents/analytic tools; K.V.N. and C.K.K. analyzed data; and K.V.N., M.S., and A.E.R. wrote the paper.

This work was supported by the National Institutes of Health (Grants R01-DC04290, R01-GM109086, and UL1-RR024979). We thank Joel Berger, Haiming Chen, Phillip Gander, Christopher Garcia, and Beau Snoad for help with data collection and analysis.

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

The authors declare no competing financial interests.

Correspondence should be addressed to Kirill V. Nourski at kirill-nourski@uiowa.edu.

https://doi.org/10.1523/JNEUROSCI.2187-21.2022

properties of other auditory-related regions, including planum temporale (PT), planum polare (PP), insula, and the anterior portion of the STG (STGA; Blenkmann et al., 2019; Woolnough et al., 2019; Zhang et al., 2019; Forseth et al., 2020; Guo et al., 2021; Hamilton et al., 2021). Most of these studies used passive listening paradigms (Blenkmann et al., 2019; Guo et al., 2021; Hamilton et al., 2021). The few studies that used active tasks restricted their examination to specific areas (Woolnough et al., 2019; Zhang et al., 2019; Forseth et al., 2020). This emphasizes the need for a comprehensive examination of the multiple auditory cortical areas within the context of active speech processing.

Efficient speech processing requires both feedforward information flow and modulatory feedback signaling from higher order areas (Rauschecker and Scott, 2009). The former is subserved by gamma band (>30 Hz) activity; the latter is indexed by lower frequency activity, particularly in the alpha (8–14 Hz) band (Bauer et al., 2014). Alpha activity, a component of feedback signaling thought to enhance auditory processing efficiency, is associated with a decrease in auditory cortical activation, and is anticorrelated with the BOLD signal (Goldman et al., 2002). Suppression of alpha band power is an important component of auditory cortical activation (Dimitrijevic et al., 2017; Billig et al., 2019; Nourski et al., 2021) and is associated with increased BOLD fMRI signal in sensory cortex (Yuan et al., 2010). Thus, measuring both gamma and alpha signals is warranted to more fully characterize speech processing.

Progression of onset latencies of gamma activation and alpha suppression across regions can clarify the position of each area within the auditory cortical hierarchy. Although HGPM, a portion of core auditory cortex (Hackett, 2015), has the earliest feedforward activation, the overall organization of auditory networks is complex and remains a matter of ongoing investigation (Humphries et al., 2014; de Heer et al., 2017; Häkkinen and Rinne, 2018). Onset latencies of alpha suppression should provide insight into the earliest targets of feedback influences from higher order areas.

Another unresolved question is the level at which hemispheric asymmetries for speech processing emerge (Leaver and Rauschecker, 2010; Hickok and Poeppel, 2015). One model posits that speech is bilaterally processed in auditory cortex, and left hemisphere bias emerges at later processing stages (Hickok and Poeppel, 2007). The other prominent model envisions a left hemisphere bias beginning early in the auditory hierarchy (Rauschecker and Scott, 2009). Additionally, to clarify whether hemispheric differences in cortical activation are speech specific or reflect more general active sound processing (Scott and McGettigan, 2013), control tasks are needed. Finally, understanding relationships between cortical activity evoked by speech and behavioral task performance will help identify key components of speech processing.

The objectives of this study were to clarify the hierarchical organization of speech processing and the emergence of hemispheric asymmetries. Gamma activation and alpha suppression were measured in a large cohort of neurosurgical epilepsy patients during a semantic classification task. A tone detection task provided a control to examine the extent to which findings could be generalized to active sound processing per se. Gamma responses recorded in HGPM, HGAL, STGP, and STGM during performance of this semantic classification task have been previously reported (Steinschneider et al., 2014; Nourski et al., 2017, 2021) for subsets of the current subject population. The present works extends those studies to include alpha suppression and the less explored areas PT, PP, insula, and STGA.

Materials and Methods

Subjects. Subjects were 29 neurosurgical patients (17 male, 12 female; age 18–52 years old; median age 34 years old) diagnosed with medically refractory epilepsy who were undergoing chronic iEEG monitoring to identify potentially resectable temporal lobe seizure foci. Demographic and iEEG electrode coverage data for each subject are presented in Extended Data Table 1-1. The hemisphere with predominant electrode coverage is indicated by the prefix of the subject code, that is, L (left, N = 15 subjects), R (right, N = 12), and B (bilateral, N = 2). Twenty-five subjects were right handed, and four were left handed. Twenty-four subjects were left-hemisphere language dominant per Wada test, one subject (R292) was right-hemisphere dominant. In the four remaining subjects (L425, R429, R456, B457), language dominance was not established, as no Wada test was performed. Of these subjects, R429, R456, and B457 were right handed, and L425 was left handed.

All subjects were native English speakers except for L275, who was a 30-year-old native Bosnian speaker who learned German at the age of 10 and English at the age of 17. Of 29 subjects, 17 had pure-tone thresholds within 25 dB hearing level (HL) between 250 Hz and 4 kHz. Subjects L258 and L425 had unilateral mild (26–40 dB HL) low-frequency (≤250 Hz) hearing loss; L372 had a unilateral 30 dB HL notch at 750 Hz; L307, R334, R369, and R376 had a unilateral mild notch at 4 kHz; L357 and L525 had unilateral mild high-frequency (≥4 kHz) loss; L282 had a bilateral mild-to-moderate notch at 2 kHz (left, 40 dB HL; right, 55 dB HL); and L423 had bilateral moderate (45–50 dB HL) high-frequency (≥4 kHz) loss. L514 was the only subject who did not undergo audiometric testing and reported normal hearing. Word recognition scores, as evaluated by spondees presented via monitored live voice, were 92/92% (left/right ear) in subject L275, 96/88% in R334, 96/92% in R369, 88/100% in R399, 92/92% in L423, 92/100% in R456, 96/92% in L525, and ≥96% in all other tested subjects. Speech reception thresholds were within 20 dB HL in all tested subjects, including those with tone audiometry thresholds outside the 25 dB HL criterion. All subjects underwent neuropsychological evaluation before the study, and none was found to have cognitive deficits that should have an impact on the findings presented in this study.

Research protocols were approved by the University of Iowa Institutional Review Board and the National Institutes of Health. Written informed consent was obtained from all subjects. Research participation did not interfere with acquisition of clinically required data, and subjects could rescind consent at any time without interrupting their clinical evaluation.

Stimuli and procedure. Experimental stimuli were monosyllabic words and complex tones presented in target detection tasks (Steinschneider et al., 2014; Nourski et al., 2017, 2021). The words were cat, dog, five, ten, red, and white, obtained from the TIMIT (Garofolo et al., 1993) and LibriVox (http://librivox.org/) databases. A total of 20 unique exemplars of each word were presented in each task, 14 spoken by different male and 6 by different female speakers. Each task also included five novel target and five novel nontarget monosyllabic words. These additional stimuli were presented to maximize the subjects' reliance on semantic cues and minimize reliance on phonemic cues. Responses to these novel words are beyond the scope of this study. The complex tones had fundamental frequencies of 125 (28 trials) and 250 Hz (12 trials), approximating fundamental frequencies of male and female talkers, respectively. Targets were either the complex tones or words belonging to the specific semantic categories of animals or numbers. The three target detection tasks (tones, animals, numbers) were presented to each subject in three separate blocks. At the beginning of each block, the subject was told by the experimenter which category the target was for that block. The subjects' task was to push a response button on a Microsoft SideWinder game controller or a USB numeric keypad when they heard a target. The subjects were instructed to use the hand ipsilateral to the hemisphere in which the majority of electrodes were implanted. This was done to reduce contributions of preparatory, motor, and somatosensory responses associated with the button press to the recorded neural activity. There was no visual component to the task, and the subjects did not receive any specific instructions other than to respond to target auditory stimuli by pressing a button.

Stimuli were delivered via insert earphones (ER4B, Etymotic Research) integrated into custom-fit ear molds. All stimuli had a duration of 300 ms and were presented in random order with an interstimulus interval chosen randomly within a Gaussian distribution (mean, 2 s; SD, 10 ms). Before each experiment, the subjects were presented with a random-sequence preview of stimuli to ensure volume was at a comfortable level (typically 55–65 dB SPL), the subjects understood task requirements, and they became acclimated to the stimulus sets. Experiments were conducted in a dedicated electrically shielded suite in the University of Iowa Clinical Research Unit. The subjects were comfortably reclining in a hospital bed or an armchair.

Recording. This study focused on iEEG recordings made from STP, the lateral STG, and the insula. Details of electrode implantation, recording and iEEG data analysis have been described previously (Nourski and Howard, 2015). In brief, depth electrode arrays (4–12 macro contacts, spaced 5 mm apart) were implanted in each subject stereotactically into HG, along its posteromedial-to-anterolateral axis. Electrodes implanted in HG provided valuable diagnostic and prognostic data regarding the efficacy of epilepsy surgery for amelioration of seizures emanating from the temporal lobe (Nagahama et al., 2018a, b). Arrays that targeted posterior and anterior insular cortex provided additional coverage of PT and PP, respectively. When used, subdural grid arrays were typically implanted over auditory cortex on the lateral STG. The grid arrays consisted of platinum iridium disk electrodes (2.3 mm exposed diameter) embedded in a silicon membrane and arranged in a 2 × 8, 4 × 8, or 8 × 12 configuration with a 5 or 10 mm center-to-center interelectrode distance.

A subgaleal contact was used as a reference in all subjects. The extent of electrode coverage varied across the subject population based on patient-specific clinically determined electrode implantation plans (Extended Data Table 1-1). Electrodes remained in place under the direction of the patients' treating neurologists.

Data acquisition was controlled by a TDT RZ2 real-time processor [Tucker-Davis Technologies (TDT)] in subjects L258 through L357 and by a Neuralynx Atlas System (Neuralynx) in subjects L369 through R532. Recorded data were amplified, filtered (0.7–800 Hz bandpass, 12 dB/octave rolloff for TDT-recorded data; 0.1–500 Hz bandpass, 5 dB/octave rolloff for Neuralynx-recorded data), digitized at a sampling rate of 2034.5 Hz (TDT) or 2000 Hz (Neuralynx), and stored for subsequent off-line analysis.

Data analysis . Anatomical locations of the implanted electrodes were reconstructed using FreeSurfer image analysis suite (https://surfer.nmr.mgh.harvard.edu) and in-house software. Subjects underwent whole-brain high-resolution T1-weighted structural MRI scans (resolution 0.78 × 0.78 mm, slice thickness 1.0 mm) before electrode implantation and again immediately after implantation. Subjects additionally underwent thin-sliced volumetric computerized tomography scans (resolution 0.51 × 0.51 mm, slice thickness 1.0 mm) postimplantation. Locations of recording sites were determined within postimplantation MR image space according to electrode-induced susceptibility artifacts and through a comparison with intraoperative photographs and coregistered postimplant computed tomography images. Contact locations were then transferred from the postoperative image space to the preoperative MRI image space using a manually guided thin-plane spline warping, for which control points were selected according to the visible correspondence between preimplant and postimplant imaging, as needed, to correct for nonlinear postoperative tissue distortion.

Locations of recording sites were projected into standard Montreal Neurologic Institute (MNI) space through a 12-parameter affine linear coregistration of the preoperative T1-weighted image with the ICBM152 sixth-generation nonlinear template brain, using the FIRST pipeline, implemented in the Functional MRI of the Brain Software Library toolbox (Jenkinson and Smith, 2001). Left-hemisphere MNI x-axis coordinates (x_{MNI}) were multiplied by (−1) to map them onto the right-hemisphere common space. Distributions of recording sites along the $|x_{MNI}|$ (medial–lateral) and y_{MNI} (posterior–anterior) axis were characterized in terms of mean MNI coordinates and their standard deviations. Comparisons between the two hemispheres and pairwise comparisons across regions of interest (ROIs) were made using two-sample two-tailed t tests. The nine ROIs examined in this study were HGPM, HGAL, PT, PP, STGP, STGM, STGA, InsP, and InsA.

Analysis of iEEG data were performed using custom software written in MATLAB version 2020a programming environment (MathWorks). Recordings were downsampled to 1000 Hz for computational efficiency and denoised using a demodulated band transform-based procedure (Kovach and Gander, 2016). Voltage deflections exceeding 5 SDs from the across-block mean for each recording site were considered artifacts, and trials containing such deflections were excluded from further analysis. To ensure that auditory cortex did not exhibit aberrant responses in any of the subjects, a series of standard protocols was performed to confirm that data from each subject conformed to the previously published corpora (Brugge et al., 2009; Nourski et al., 2013, 2014, 2015, 2016; Nourski, 2017; Steinschneider et al., 2011).

Time-frequency analysis was implemented using the demodulated band transform-based algorithm (Kovach and Gander, 2016). This was done by computing the discrete Fourier transform over the entire duration of the recording, segmenting the discrete Fourier transform into overlapping windows of 1, 2, 4, 10, and 20 Hz bandwidth for theta (4–8 Hz), alpha (8–14 Hz), beta (14–30 Hz), low gamma (30–70 Hz), and high gamma (70–150 Hz) bands, respectively. Event-related band power (ERBP) was calculated by log transforming power for each center frequency and normalizing it to a baseline value measured as the mean power in the prestimulus reference interval (100–200 ms before stimulus onset).

ERBP spectral profiles were examined by averaging ERBP values for each center frequency between 4 and 150 Hz within 50–350 ms after stimulus onset, as done previously by Nourski et al. (2021). To determine gamma iEEG frequencies associated with largest ERBP values, time-frequency analysis of responses to words in all target detection tasks (complex tone, animal, and number target conditions) was conducted for center frequencies from 30 to 150 Hz in 1 Hz steps. For each recording site, ERBP values for each center frequency were averaged within the 50–350 ms poststimulus window. A concave quadratic function was fit to the iEEG frequency versus ERBP data using the MATLAB polyfit function. Gamma ERBP peak frequencies were measured as the peak of the quadratic fit at sites where the leading coefficient of the fit was greater than −0.0001, as less negative coefficients were indicative of an absence of a discernable peak in the gamma ERBP spectral profile.

For quantitative analyses, gamma ERBP and alpha ERBP were averaged within the 50–350 ms and 350–650 ms poststimulus windows, respectively (Nourski et al., 2021). Stimulus effect was defined as the difference in ERBP between responses to nontarget words and complex tones. Target effect was defined as ERBP difference between responses to nontarget and target words. Gamma ERBP onset latencies were defined as the time at which the lower limit of gamma ERBP 95% confidence interval exceeded 0 dB relative to the prestimulus mean and remained positive for at least 30 ms, following the approach of Nourski et al. (2014, 2021). Onset latencies of alpha ERBP suppression were defined as the time at which the upper limit of alpha ERBP 95% confidence interval was below 0 dB relative to the prestimulus mean and remained negative for at least 120 ms, as done previously in Nourski et al. (2021).

Linear mixed-effects models were used to compare gamma peak frequencies, gamma and alpha ERBP, and gamma and alpha onset latency (Output) between ROIs. The fixed effect of ROI was modeled as a nine-level contrast for ERBP (i.e., including all ROIs), and as a seven-level contrast (all ROIs except STGA and InsA) for gamma peak frequencies and gamma and alpha response latency. STGA and InsA were not included in these latter analyses because of the relatively small (<10) number of sites in these ROIs that had large enough responses to reliably measure the response properties using the approach summarized above. All models included a fixed effect of hemisphere and a random intercept for subject. The following models were implemented in MATLAB version 2020a using the fitlme function as follows:

$$\text{Output} \sim \text{ROI} * \text{Hemisphere} + (1 \mid \text{Subject}).$$

A set of complementary analyses examined gradients of gamma activation and alpha suppression (Output) with respect to location of recording sites within MNI coordinate space in the three anatomic planes (STP, STG, and insula). MNI coordinates were transformed (centered and rotated) to project the three-dimensional location of a recording site into a two-dimensional space aligned with the first two principal axes of the corresponding anatomic structure (STP, x_θ and y_θ; STG and insula, y_θ and z_θ, respectively). To accomplish this, a rigid affine transformation was obtained from a singular value decomposition (SVD) applied to selected vertices of the pial surface mesh generated by FreeSurfer with the 2006 ICBM MNI-152 template brain. Vertices were selected according to the label assigned within the Desikan–Killiany–Tourville atlas parcellation (Klein and Tourville, 2012). In each case, SVD was applied to the selected vertex coordinates after subtracting the center of mass, and the first two dimensions of the resulting affine transform were retained for the purpose of linear mixed-effects modeling.

For STP, the models included main effects of location along and across HG (x_θ and y_θ, respectively), a quadratic term describing a curvilinear relationship on either side of the crest of HG ($y_\theta{}^2$), their interactions with hemisphere, and a random intercept for subject as follows:

$$\text{Output} \sim x_\theta * y_\theta{}^2 * \text{Hemisphere} + (1 \mid \text{Subject}).$$

For STG, the models included main effects of location along and across STG (y_θ and z_θ, respectively), a quadratic term describing a curvilinear relationship along the posterior–anterior axis of the STG ($y_\theta{}^2$; Nourski et al., 2014), their interactions with hemisphere, and a random intercept for subject as follows:

$$\text{Output} \sim y_\theta{}^2 * z_\theta * \text{Hemisphere} + (1 \mid \text{Subject}).$$

For insula, the models included main effects of location along the posterior–anterior and ventral–dorsal axis of insular cortex (y_θ and z_θ, respectively), and a random intercept for subject as follows:

$$\text{Output} \sim y_\theta * z_\theta + (1 \mid \text{Subject}).$$

Models that examined stimulus and target effects (Condition) on ERBP magnitude (Output) across recording sites included a random effect for stimulus and random intercept of recording site (Channel) nested within subject as follows:

$$\text{Output} \sim \text{Condition} * \text{Hemisphere} + (1 \mid \text{Subject} : \text{Channel}) + (1 \mid \text{Subject}).$$

The fixed effects of stimulus and target were modeled as two-level contrasts (stimulus effect, nontarget tones vs nontarget words; word target effect, nontarget words vs target words; tone target effect, nontarget tones vs target tones). In all model-based analyses, subject was included as a random effect to minimize the possibility that significant fixed effects were biased by single-subject contributions. Hemispheric asymmetries were examined by adding hemisphere as a fixed effect to the models and comparing the model fit (with and without main effect of Hemisphere) using theoretical likelihood ratio tests. The quality of models was evaluated using Akaike Information Criterion (AIC).

Statistical significance of differences in gamma and alpha ERBP between subjects who had above-average (>82% hit rate; $N = 15$) and below-average (<82% hit rate; $N = 14$) performance in the semantic categorization tasks was established using Wilcoxon rank sum tests. To correct for multiple comparisons, positive false discovery rate (FDR) for multiple hypothesis testing was controlled using the approach of Benjamini and Hochberg (1995), and p values were adjusted as the measures of hypothesis testing error.

Results

Electrode coverage

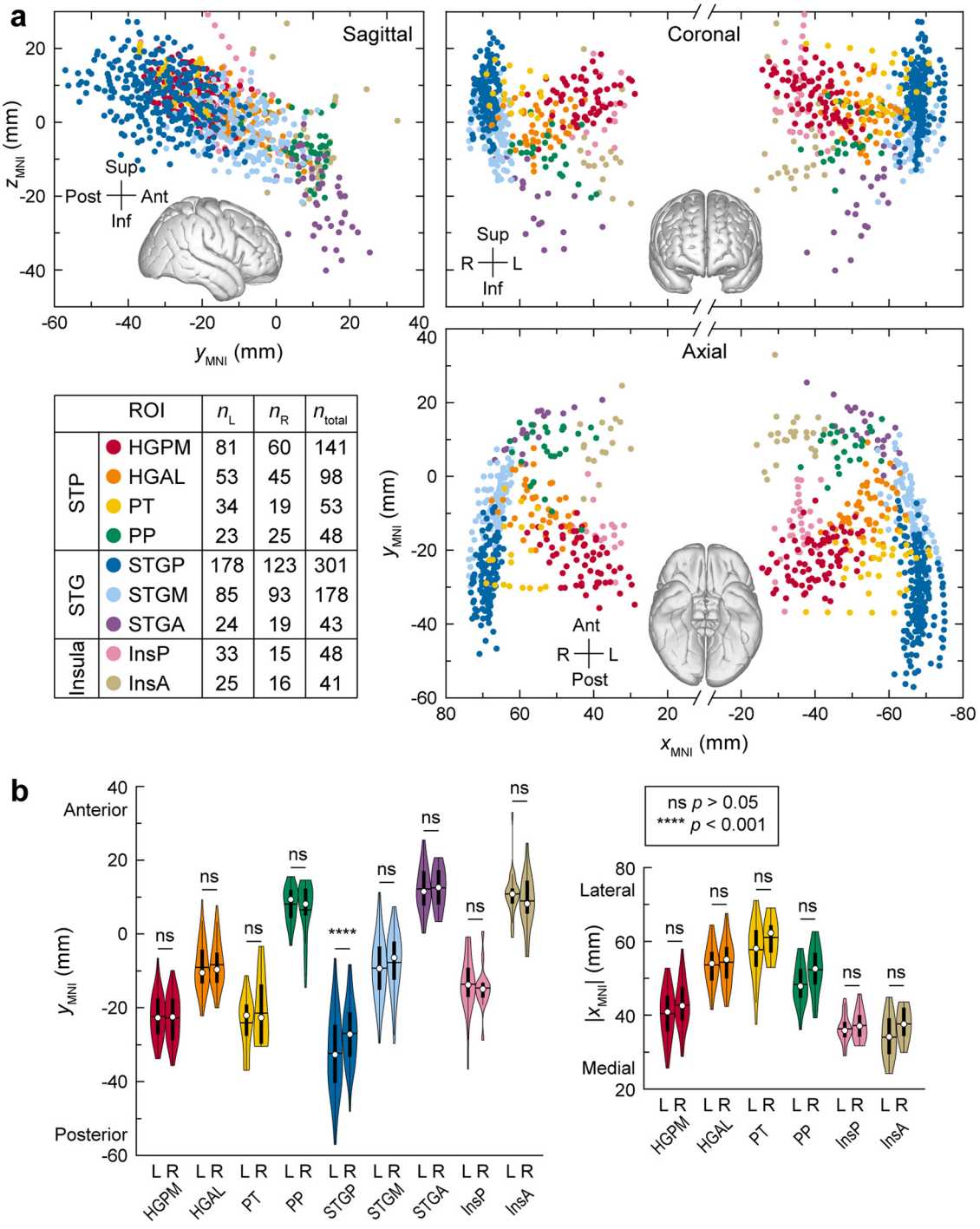
Results are based on analysis of data from 29 subjects who had electrode coverage of the STP and, to a variable extent, from the lateral STG and insula (Fig. 1; Extended Data Table 1-1). The distribution of recording sites in all nine ROIs across all subjects is plotted in standard MNI coordinate space in Figure 1*a*. All 29 subjects had depth electrodes within core auditory cortex (HGPM, 141 sites) and 24 subjects had coverage of adjacent noncore auditory cortex of HGAL (98 sites). Fifteen subjects had coverage of PT (53 sites), and 16 subjects had recording sites in PP (48 sites). Subdural grid arrays provided coverage of the lateral STG in 26 subjects (301, 178, and 43 sites in the posterior, middle, and anterior portion of the gyrus, respectively). Nineteen subjects had coverage of InsP (48 sites), and 17 subjects contributed a total of 41 sites in InsA. This sample size compares favorably with those in recent iEEG studies that examined subsets of these ROIs (Blenkmann et al., 2019; Woolnough et al., 2019; Zhang et al., 2019; Forseth et al., 2020; Guo et al., 2021; Hamilton et al., 2021).

Figure 1*b* depicts distributions of sites along the y_{MNI} (posterior to anterior) and $|x_{MNI}|$ (medial to lateral) axis. There was no systematic difference in the spatial extent of coverage between the two hemispheres for all ROIs except for STGP, where coverage in the left hemisphere extended more posteriorly than in the right ($p < 0.0001$; Fig. 1*b*). This difference likely reflects the known hemispheric asymmetry in the human temporal lobe anatomy, wherein left superior temporal gyrus extends more posteriorly than its right hemisphere counterpart (Geschwind and Levitsky, 1968; Steinmetz et al., 1990; Loftus et al., 1993). Thus, comparable spatial extent of coverage of both the left and right hemispheres permitted comparisons of physiologic response properties of each ROI between the two hemispheres.

Spectral characteristics of responses to words

Responses elicited by monosyllabic words within auditory cortex have characteristic spectrotemporal profiles. These profiles are depicted for each ROI separately in the grand average time-frequency plots illustrated in Figure 2*a*. The initial response in all ROIs, to varying degrees, consisted of increases in both low (γ_{low}; 30–70 Hz) and high (γ_{high}; 70–150 Hz) gamma ERBP that were maximal within the first several hundred milliseconds after stimulus onset. These initial increases were followed by suppression of lower frequency power centered in the alpha band (alpha, 8–14 Hz). The time-frequency windows used for subsequent quantitative analyses, which encompassed much of the changes in gamma and alpha ERBP, are denoted by dashed rectangles in Figure 2*a* (gamma, 30–150 Hz, 50–350 ms; alpha, 8–14 Hz, 350–650 ms). Increases in gamma ERBP and alpha suppression within PP, STGA, and InsA were weaker than in the other ROIs, likely a manifestation of their more downstream locations in the auditory cortical hierarchy.

Locations outside of core areas can exhibit different degrees of event-related power increases within low and high gamma iEEG frequency bands (Mahjoory et al., 2020; Nourski et al., 2021). These differences have been suggested to reflect differences in cytoarchitectonic organization among the various auditory cortical regions (Mahjoory et al., 2020). Accordingly, ERBP spectral profiles of responses to the monosyllabic words were examined in the early (50–350 ms) analysis window, when both low



	ROI	n_L	n_R	n_{total}
STP	HGPM	81	60	141
	HGAL	53	45	98
	PT	34	19	53
	PP	23	25	48
STG	STGP	178	123	301
	STGM	85	93	178
	STGA	24	19	43
Insula	InsP	33	15	48
	InsA	25	16	41

Figure 1. Electrode coverage of STP, STG, and insula in all 29 subjects. ***a***, Locations of recording sites, color coded according to the ROI, are projected onto sagittal, coronal, and axial planes in MNI coordinate space. Lateral, front, and bottom-up views of the average FreeSurfer brain are shown for reference. ***b***, Distribution of all recording sites in the left (L) and right (R) hemisphere along posterior–anterior axis (y_{MNI}) and sites in STP and insula along the medial–lateral axis ($|x_{MNI}|$; left and right, respectively). In each violin plot, white circle denotes the median, horizontal line denotes the mean, bar denotes Q_1 and Q_3, and whiskers show the range of lower and higher adjacent values (i.e., values within 1.5 interquartile ranges below Q_1 or above Q_3, respectively). Extended Data Table 1-1 shows subject demographics and electrode coverage.

and high gamma augmentation were greatest (Fig. 2*b*). Peak frequencies were greatest in core auditory cortex located in HGPM (99 ± 16 Hz; mean± SD; Fig. 2*c*). Gamma ERBP peak frequencies within HGPM were significantly higher than in all other ROIs ($p = 0.017$ for InsP vs HGPM $p < 0.001$ for all other contrasts; Extended Data Table 2-1 contains details of linear mixed-effects model analysis). The three regions with the lowest peak frequencies (PP, HGAL, and PT, the latter having the lowest mean peak frequency of 84 ± 20 Hz) were all located immediately adjacent to HGPM. Thus, to account for this gamma frequency shift across auditory cortical regions, low gamma and high gamma power were combined into a single event-related band (30–150 Hz) for all subsequent analyses. Gamma responses in STGA and InsA sites were too weak to allow for reliable measurements of their peak frequencies.

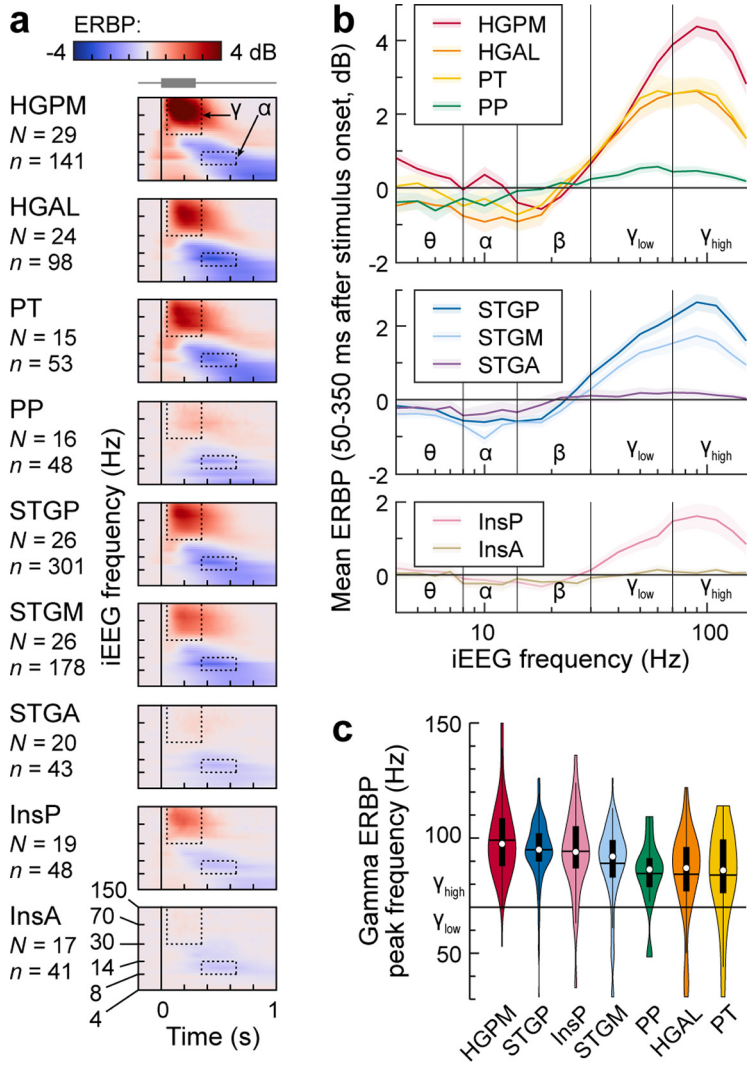


Figure 2. Activity in auditory and auditory-related cortex during auditory target detection tasks. ***a***, Time frequency plots of responses to the monosyllabic words cat, dog, five, ten, red and white, averaged across all recording sites in each of the nine ROIs (top to bottom). Each word was spoken by 20 talkers (14 male and 6 female). Top, Stimulus schematic (duration 300 ms) in gray. Dotted rectangles denote the time windows used for quantitative analyses of gamma (γ; 50–350 ms, 30–150 Hz) and alpha (α; 350–650 ms, 8–14 Hz) ERBP. ***b***, ERBP spectral profiles of responses in the nine ROIs. ERBP values, averaged within the 50–350 ms poststimulus time window, are plotted as functions of iEEG frequency for each ROI (STP, STG, and insula ROIs, top, middle, and bottom, respectively). Colored lines and shading represent across-site means and their 95% confidence intervals. ***c***, Distributions of gamma ERBP peak frequencies across sites; ROIs rank ordered by mean gamma peak frequency value. In each violin plot, white circle denotes the median, horizontal line denotes the mean, bar denotes Q_1 and Q_3, and whiskers show the range of lower and higher adjacent values (i.e., values within 1.5 interquartile ranges below Q_1 or above Q_3, respectively). Note that for iEEG frequencies, logarithmic scale is used for the ordinate in ***a*** and abscissa in ***b***, and linear scale is used for the ordinate of ***c***. Data from STGA and InsA are not shown, as weak responses in these ROIs precluded reliable measurements of peak gamma frequencies. Extended Data Table 2-1 shows results of model-based pairwise comparisons between ROIs.

Differential profiles of gamma augmentation and alpha suppression across the STP, STG, and insula

To examine profiles of gamma augmentation and alpha suppression across the STP, STG, and insula, analyses were performed using models with ROI and hemisphere as fixed effects with random intercept for subject (Fig. 3*a*; Extended Data Table 3-1). Largest gamma responses were observed in HGPM ($p < 0.0001$ for pairwise comparisons with all other ROIs). Activity in PT, HGAL, and STGP was comparable in magnitude and was larger than in the remaining ROIs. Finally, STGM and InsP had comparable responses that were larger than those in PP, STGA, and InsA. Hemispheric asymmetries for gamma ERBP were limited to STGP, with larger responses in the right hemisphere ($p = 0.00012$).

Gamma activation was earliest in HGPM, followed by PT and InsP, which in turn had shorter onset latencies than HGAL, PP, STGP, and STGM (Fig. 3*b*; Extended Data Table 3-2). Rank ordering ROIs based on their gamma onset latencies revealed a pattern wherein caudal ROIs (PT, InsP, STGP) had shorter latencies than more rostral ROIs (HGAL, STGM, PP). Onset latencies in the right hemisphere were significantly shorter in STGP, STGM, and PP ($p = 0.0016$, $p = 0.033$, $p = 0.012$, respectively). This spatial organization of gamma ERBP and onset latency was confirmed using a complementary approach with respect to location of recording sites within MNI coordinate space (Extended Data Fig. 3-1, Extended Data Tables 3-3, 3-4).

The regional distribution of alpha suppression was markedly different from that seen for gamma enhancement. Most notably, alpha suppression was maximal in HGAL, followed by PT and STGP (Fig. 3*c*; Extended Data Table 3-5). It was also more pronounced in the left hemisphere in HGAL, PP, and all three subdivisions of STG. These findings, including greater alpha suppression in the left STP, were mirrored by model analyses of the three anatomic planes (Extended Data Fig. 3-2*a*; Extended Data Table 3-6). Additionally, there was stronger alpha suppression at STG locations near the Sylvian fissure compared with those near superior temporal sulcus. Alpha suppression was uniformly weak in the insula.

Latency of alpha suppression also demonstrated a markedly different profile from that seen for latency of gamma responses. The onset latency of alpha suppression in HGPM was longer than that seen in all other ROIs except InsP (Fig. 3*d*; Extended Data Table 3-7), a reverse pattern than that seen for gamma enhancement. Latencies were shortest in STGM, followed by HGAL, STGP, and PP, and longer in other ROIs. Additionally, STGM was characterized by shorter onset latencies of alpha suppression in the left hemisphere ($p = 0.016$), contrasting with shorter gamma response latencies in the right hemisphere. Alpha suppression onset latency gradients paralleled those seen for alpha power at a regional level (Extended Data Fig. 3-2*b*; Extended Data Table 3-8).

Thus, gamma augmentation and alpha suppression had different profiles across the STP, STG, and insula, supporting the interpretation that gamma and alpha iEEG power changes represent two distinct neural processes involved in sound processing. The former reflects mechanisms involved in feedforward sound processing, whereas the latter reflects release of top-down

inhibition from higher cognitive areas during sound processing (Chao et al., 2018).

Stimulus and target effects in the semantic categorization task

The semantic categorization task permitted the examination of stimulus effect, that is, differences in response magnitude between nontarget complex tones and nontarget words, and target effect, which reflects additional processing that ultimately leads to correct identification of target stimuli. The tone detection task provided a control to establish whether the target effect in the semantic categorization task was speech specific or a more general effect of active sound processing. Nontarget trials that were associated with incorrect responses and target trials to which subjects failed to respond were excluded from this analysis.

Stimulus effects were examined by measuring gamma activity and alpha suppression in the same two successive time windows used in previous analyses (50–350 and 350–650 ms after stimulus onset, respectively; Fig. 4*a*). Target effect was solely measured using alpha suppression, as previous work using the same paradigm had shown that gamma activity reflecting target effect occurred primarily (>90%) in higher order cortical areas (e.g., MTG, SMG, and IFG in the left hemisphere; Nourski et al., 2017).

Gamma stimulus effect was present in all ROIs except for the three areas with the weakest responses (PP, STGA, InsA; Fig. 4*b*; Extended Data Table 4-1). There was a stronger gamma stimulus effect (i.e., larger responses to words vs tones) in HGAL, STGP, and STGM in the left hemisphere. The analysis of this effect was extended by comparing gamma responses to tones and words on a site-by-site basis (Fig. 5). Sites with stronger responses to words than tones were more prevalent than sites that exhibited the opposite relationship, as could be expected from the ROI-level analysis depicted in Figure 4*b*. The opposite pattern of stronger responses to tones versus words occurred more frequently in the right hemisphere than the left. In the left hemisphere, 17 of 536 sites were characterized by significantly stronger ($p < 0.05$, two-sample t tests, FDR corrected) gamma responses to tones than words; 14 of these were in STGP, 2 in HGPM, and 1 in HGAL. By contrast, stronger responses to tones were found in 55 of 415 right-hemisphere sites. This effect was noted in all ROIs except InsP. For STGP, there was a right-hemisphere bias for the complex tones to elicit larger gamma responses compared with words (right hemisphere, 31 of 123 sites; left hemisphere, 14 of 178 sites; χ^2 test $p < 0.0001$). Thus, although gamma stimulus effect typically manifested as elevated responses to words compared with tones, sites within the right hemisphere, especially STGP, could exhibit the opposite effect. Of note, the two complex tones were constructed to mimic vowel-like sounds with fundamental frequencies typical of male and female voices. This suggests that although both hemispheres are active in processing speech, the right hemisphere may be focused on processing specific components of the speech stream (e.g., vowels and pitch).

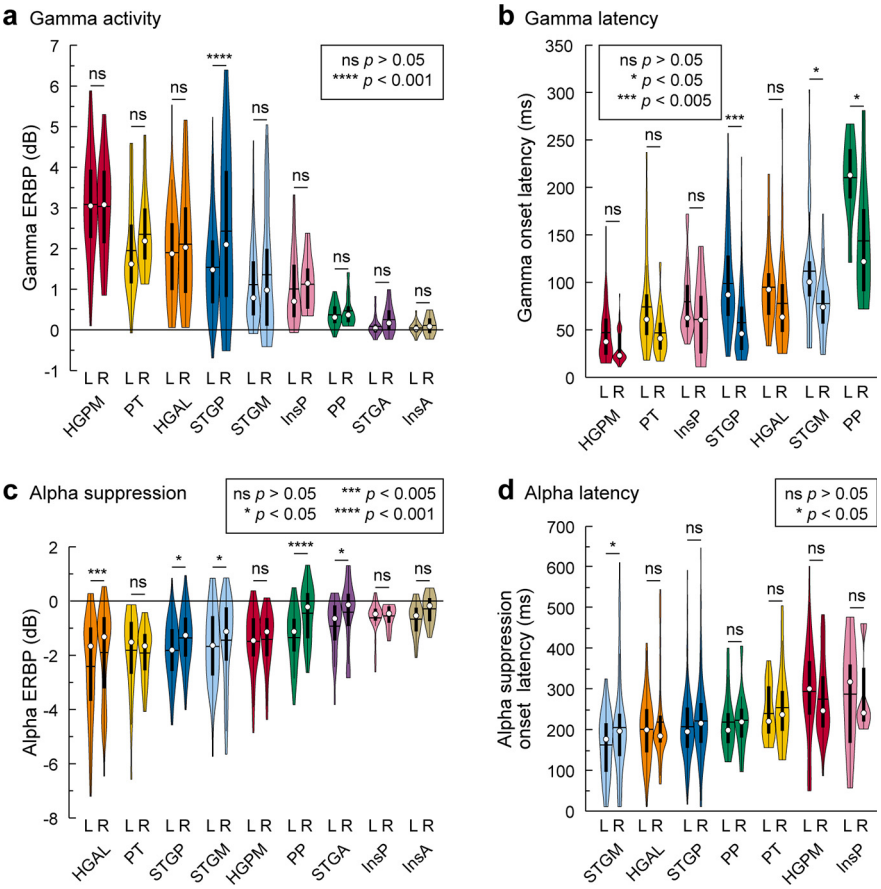


Figure 3. Gamma activity and alpha suppression in auditory and auditory-related cortex during the semantic categorization task. Data are plotted by ROI and hemisphere, left (L), right (R). ***a***, Distribution of gamma ERBP values across sites. ROIs are rank ordered by mean gamma ERBP value. In each violin plot, white circle denotes the median, horizontal line denotes the mean, bar denotes Q_1 and Q_3, and whiskers show the range of lower and higher adjacent values (i.e., values within 1.5 interquartile ranges below Q_1 or above Q_3, respectively). ***b***, Distribution of gamma onset latency values across sites; ROIs rank ordered by mean gamma onset latency. Data from STGA and InsA are not shown, as weak responses in these ROIs precluded reliable measurements of gamma onset latencies. ***c***, Distribution of alpha ERBP values across sites. ROIs are rank ordered by mean alpha ERBP value. ***d***, Distribution of alpha suppression onset latency values across sites. ROIs rank ordered by mean alpha suppression onset latency. Data from STGA and InsA are not shown, as weak responses in these ROIs precluded reliable measurements of alpha onset latencies. Extended Data Tables 3-1, 3-2, 3-5, and 3-6 show results of model descriptions and quantitative results of pairwise comparisons between ROIs.

Alpha stimulus effect (greater alpha suppression in response to words vs tones) was present in all ROIs except for the insula (Fig. 4*c*; Extended Data Table 4-2). At the single-site level, alpha stimulus effect was much weaker than gamma stimulus effect (27 vs 310 of 951 sites, for alpha and gamma, respectively). InsP was different from canonical auditory cortex in that it exhibited a gamma/alpha dissociation, wherein words elicited larger early gamma activation than tones without a subsequent modulation of alpha power. A larger alpha stimulus effect was noted in the left hemisphere in HGPM, STGP, and STGM and paralleled the more general left-hemispheric bias of the alpha suppression (Fig. 3*c*).

Alpha target effect in the semantic categorization task (alpha word target effect, i.e., greater alpha suppression in response to

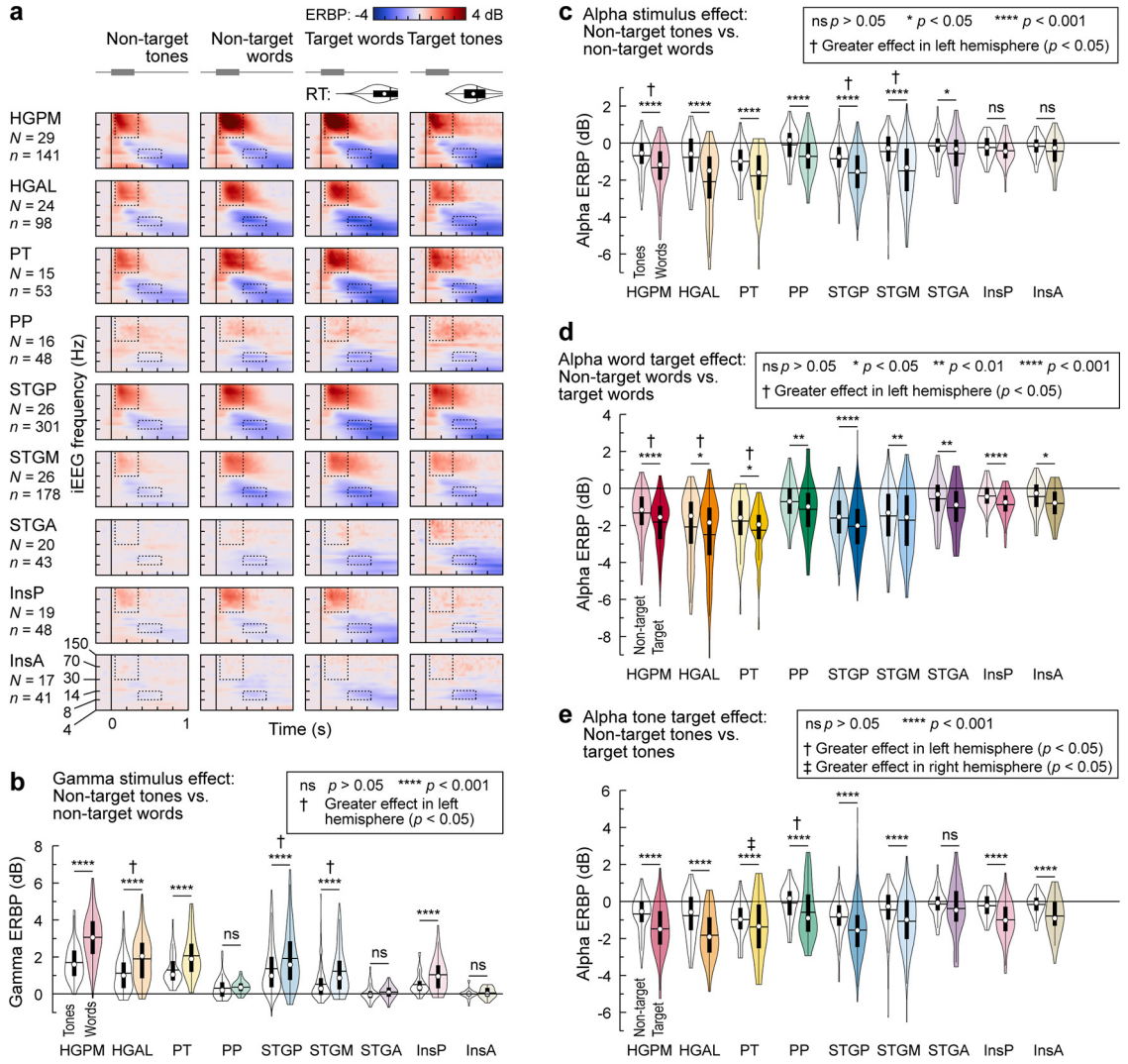


Figure 4. Stimulus and target effects in auditory target detection tasks. ***a***, Time-frequency plots of responses to complex tones, presented as nontarget and target stimuli (first and fourth column, respectively) and monosyllabic words cat, dog, five, and ten, presented as nontarget and target stimuli (second and third column, respectively). Responses were averaged across all recording sites in each of the nine ROIs (top to bottom). Responses were only included in the averages if they were correct rejections for nontarget trials and correct hits for target trials. Top, Stimulus schematics are in gray, and the distribution of reaction times (RTs) for target hits is shown above responses to the target trials. Dotted rectangles denote the time windows used for quantitative analyses of gamma (50–350 ms, 30–150 Hz) and alpha ERBP (350–650 ms, 8–14 Hz). ***b***, Gamma stimulus effect. Distributions of gamma ERBP across sites in each ROI, grouped by condition (nontarget tones vs nontarget words). ***c***, Alpha stimulus effect. Distributions of alpha ERBP across sites in each ROI, grouped by condition (nontarget tones vs nontarget words). ***d***, Alpha word target effect. Distributions of alpha ERBP across sites in each ROI, grouped by condition (nontarget words vs target words). ***e***, Alpha tone target effect. Distributions of alpha ERBP across sites in each ROI, grouped by condition (nontarget tones vs target tones). Statistical significance of differences was established using models with stimulus condition (nontarget tone, nontarget word, target word, or target tone) and hemisphere as fixed effects and subject as the random effect. The symbols † and ‡ denote significant interactions with hemisphere (greater effects in the left and right hemisphere, respectively). Extended Data Tables 4-1 and 4-2 contain a description of models and results.

target vs nontarget words) was observed in all ROIs, including those where responses were of generally low magnitude (PP, InsA, and STGA) and where analysis failed to reveal an alpha stimulus effect (InsP and InsA; Fig. 4*d*, Extended Data Table 4-2). A left-hemispheric bias for this effect was observed in HGPM, HGAL, and PT. This alpha target effect could be a general property of active sound processing or be specific for the processing of speech. To disambiguate these possibilities, we examined alpha target effect in the tone detection task by comparing alpha suppression associated with responses to nontarget versus target tones (Fig. 4*e*). Target tones elicited a greater degree of alpha suppression in all ROIs except STGA, indicating that the alpha target effect is a general property of active sound processing and is not specific to speech. There were, however, hemispheric asymmetries in PT and PP. In PT, there was a left-hemispheric bias for target over nontarget words and a right-hemispheric bias for tone targets relative to nontarget tones. In PP, where the word target effect was comparable between the two hemispheres, a left-hemispheric bias emerged for tone targets. Thus, data indicate that alpha target effects represent a fundamental property of active sound processing and exhibit hemispheric asymmetries, biased toward left hemisphere in HGPM, HGAL, and PT for words, and biased toward the right PT and the left PP for tones.

Relationships between physiology and behavioral performance

Relationships between physiology in sensory areas and behavioral performance are not straightforward because multiple processing stages beyond sensory cortex contribute to task completion. However, analysis of activity in sensory cortex may indicate key components that help shape the ultimate behavior. This question was examined by comparing responses to all words in the

semantic categorization task between subjects with above- versus below-average task performance (with 82% hit rate as threshold; Fig. 6). A similar analysis could not be performed for the tone detection task because of a ceiling effect in task performance. Subjects performed very well on this task, with 24 of 29 having had a hit rate of 90% or better. Median hit rate in this task was 97.5% (i.e., one missed target stimulus of 80), compared with median hit rate of 82% in the semantic categorization task.

In the semantic categorization task, the only ROI where gamma ERBP was significantly larger in subjects who did better on the task was PT (p = 0.031). This effect was restricted to the left hemisphere (p = 0.00053). Greater alpha suppression was associated with better task performance in HGPM and HGAL (p = 0.00034 and p = 0.00025, respectively), and this relationship was present in both hemispheres.

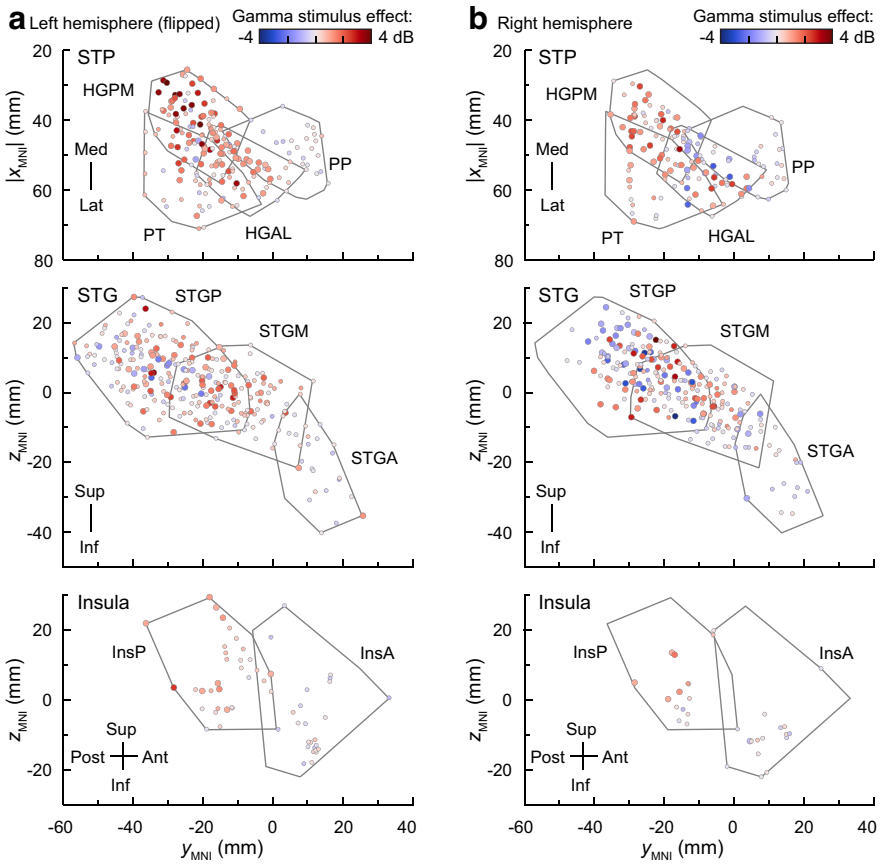


Figure 5. Spatial distribution of gamma stimulus effect in the left ***(a)*** and right ***(b)*** hemisphere sites within STP, STG and insula (top, middle and bottom plots). Positive values (shades of red) represent larger responses to non-target words vs. tones; negative values (shades of blue) represent larger responses to tones vs. non-target words. Larger and smaller circles represent significant and non-significant differences, respectively, on an individual site level.

Discussion

Summary of findings

This study used iEEG to characterize active speech processing within human STP, lateral STG, and insula in a large subject cohort. Four major findings are reported. The first reflects the complementary roles of broadband gamma augmentation and alpha suppression. The second promotes a framework for the hierarchical organization of auditory cortex based on onset latencies of both gamma augmentation and alpha suppression. The third identifies the emergence of hemispheric asymmetries early in the speech processing stream. Finally, gamma augmentation and alpha suppression can be directly related to performance in the semantic categorization task.

Spectrotemporal properties of responses to words

Responses to monosyllabic words were characterized by two major components, an earlier gamma power increase and a later suppression of alpha power. Previous iEEG studies of auditory cortical processing have typically focused on high gamma activity (Crone et al., 2001; Steinschneider et al., 2008; Nourski et al., 2009; Mesgarani et al., 2014). The present study revealed that gamma activation spanned both low and high gamma bands, with a downward shift of peak frequencies in ROIs adjacent to HGPM. Thus, the two gamma bands were combined into a single measure. ERBP is calculated for each frequency band by normalizing power to a baseline within that band. Therefore, results are not biased by the 1/f spectral property of the electroencephalogram. Auditory-related cortex within the superior temporal sulcus shares a similar downward gamma frequency shift (Nourski et al., 2021). Cytoarchitectonic differences across fields may account for the changes in spectral profiles (Morel et al., 1993; Brunel and Wang, 2003; Traub et al., 2003).

Suppression of alpha power is generally interpreted as a decrease in top-down modulation (Billig et al., 2019). Alpha suppression may represent a means by which neuronal ensembles undergo a transition from a lower to a higher excitability state and desynchronize coactivation, facilitating processing of relevant stimulus attributes (Klimesch et al., 2007; Jensen and Mazaheri, 2010; Billig et al., 2019). Alpha suppression provides a complementary measure to gamma augmentation when studying auditory cortical organization, functional asymmetries, and relationships between physiology and behavioral performance. Alpha suppression operates over relatively extensive areas, whereas gamma excitation occurs in more discrete neuronal populations (Crone et al., 2011), which may be too small to be reliably identified on the mesoscopic scale offered by iEEG. Alpha suppression can be reliably measured noninvasively by electroencephalography and magnetoencephalography and thus may have translational relevance when studying normal and aberrant speech processing (Başar and Güntekin, 2012).

Hierarchical organization of auditory cortex

Cortical gamma activity reflects feedforward signaling, and regional distribution of onset latencies can provide an opportunity to build a tentative hierarchical model of auditory cortex. HGPM

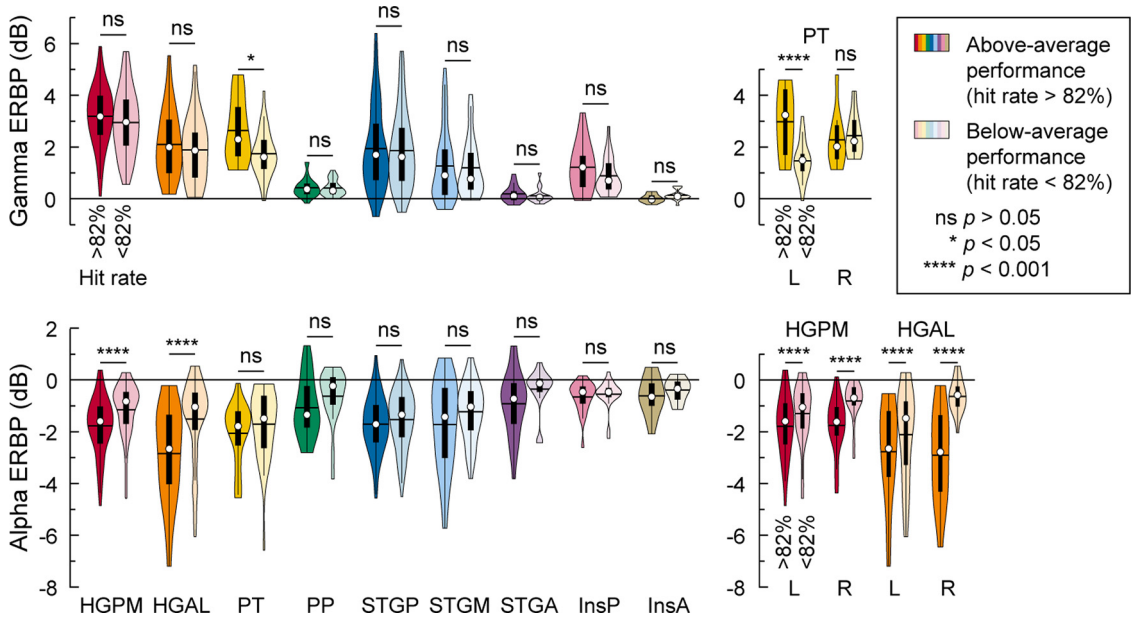


Figure 6. Relationship between iEEG responses to word stimuli and performance in the semantic categorization task. Distribution of gamma and alpha ERBP (top and bottom, respectively) in response to monosyllabic words across sites in all ROIs, plotted separately for subjects with above-average (target hit rate >82%; $N = 15$, darker violin plot in each pair) and below-average (target hit rate <82%; $N = 14$, lighter violin plot in each pair) performance in the semantic categorization task. Right, The distributions separately for right and right hemisphere for ROIs that exhibited significant differences in ERBP between the two groups of subjects. Statistical significance of differences was established using Wilcoxon rank sum tests.

has the strongest and earliest gamma responses, as expected from a core sensory area. PT has the second-shortest onset latencies. Like HGPM, PT encodes absolute pitch, whereas HGAL represents the more complex attribute of relative pitch (Hamilton et al., 2021). Combined, these results suggest that PT is an earlier processing stage than HGAL. The short latency responses in PT are reminiscent of those seen in caudal belt areas in the monkey and may reflect direct inputs from the auditory thalamus (Burton and Jones, 1976; Camalier et al., 2012).

Anatomical division of lateral STG into STGP and STGM along the transverse temporal sulcus (Upadhyay et al., 2008) is paralleled by differences in response profiles between these ROIs, including shorter onset latencies of gamma activity in STGP (Hamilton et al., 2018, 2021), thus placing it earlier in the hierarchy than STGM. Onset latencies in HGAL are shorter than in STGM, placing STGM later in the hierarchy compared with both STGP and HGAL. The shortest alpha suppression onset latencies occurred in the STGM and HGAL, suggesting that these regions are the initial targets of attenuated top-down influences in auditory cortex. Given that alpha activity from higher order regions would most likely first engage higher levels of auditory cortex, a reciprocal relationship can be expected between the onset of gamma activation and alpha suppression. This is indeed the case when considering functional relationships of STGM and HGAL with HGPM and PT.

InsP also represents an early stage in auditory cortical processing. Infarction in this region can result in auditory agnosia (Habib et al., 1995; Bamiou et al., 2003), highlighting its importance for auditory processing. Like HGPM, InsP is characterized by responses with short onset latencies and can track the fundamental frequency of male voices (Zhang et al., 2019). This isomorphic representation is a signature of early cortical sound processing (Nourski and Brugge, 2011; Steinschneider et al., 2013) and may reflect direct inputs to InsP from the auditory thalamus (Burton and Jones, 1976).

The rostral areas PP, STGA, and InsA exhibited weak gamma responses during performance of the word classification task. This weak activation may reflect functional differences of these areas with respect to the other ROIs. Functional neuroimaging has shown strong activation of PP and STGA in response to music (Leaver and Rauschecker, 2010; Angulo-Perkins et al., 2014). With respect to speech, PP and STGA are envisioned to play a role in sentence-level processing (Friederici et al., 2000; Humphries et al., 2005) rather than the task of semantic analysis of single words. Finally, InsA is strongly activated by the emotional content of speech (Zhang et al., 2019), which was not explored by the current experimental design.

Hemispheric asymmetries

Hemispheric asymmetries of gamma activity emerged in the noncore auditory areas HGAL, STGP, and STGM, wherein enhanced responses to words versus complex tones (stimulus effect) were more prominent in the left hemisphere. In contrast, more sites in the right STGP had larger responses to tones than words. At present, it is unclear whether these asymmetries are determined by acoustic complexity or speech-related mechanisms. In support of the latter interpretation, it has been proposed that the left hemisphere has a greater role in processing rapidly changing components of speech (e.g., consonants), whereas the right hemisphere is specialized for tracking slower features (e.g., vowels; Zatorre et al., 2002; Zatorre and Gandour, 2008), which are approximated by the complex tones. The strong pitch strength of these sounds may also be a factor contributing to the larger number of sites activated by tones versus words in the right STGP (Boemio et al., 2005).

Alpha target effects were biased toward the left hemisphere in HGPM, HGAL, and PT for words, and toward the right PT and the left PP for tones. Scalp-recorded event-related potentials elicited in a similar semantic classification task localized word target effects to the left temporoparietal cortex (Shahin et al., 2006). It is reasonable to speculate that the asymmetries identified in the current iEEG study contribute to those seen at the scalp. Stronger alpha tone target effect in the right PT parallels asymmetric pitch processing in this field (Hyde et al., 2008). It is too early to speculate on the lateralization of alpha target effect for tones in PP.

Cortical responses and behavioral performance

Increased alpha suppression in HGPM and HGAL bilaterally was associated with better task performance. Removal of a top-down inhibition via alpha suppression may facilitate feedforward processing and ultimately lead to better task performance. Bilateral involvement of HGPM and HGAL and its association with task performance supports models that envision the importance of bilateral processing of speech early in the auditory cortical hierarchy (Hickok and Poeppel, 2007; 2015; Leff et al., 2008; Hickok, 2009).

A possible contradiction arises when considering the PT—an early stage in the auditory hierarchy—where gamma activity in the left, but not the right hemisphere, was associated with better task performance. Functional neuroimaging data also demonstrate the importance of the left PT for phonological processing (Vouloumanos et al., 2001; Griffiths and Warren, 2002; Jäncke et al., 2002). Thus, the current data support, in part, both major models of speech processing (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009) with regard to early versus late emergence of hemispheric asymmetries and argue for a more nuanced interpretation that incorporates features of both models.

Caveats and future directions

Caution must be exercised when interpreting weak activation in the three rostral areas (STGA, PP, and InsA). These areas are engaged in multiple auditory-related functions including sentence-level processing, music listening, and perception of emotional content of speech, respectively (Friederici et al., 2000; Leaver and Rauschecker, 2010; Angulo-Perkins et al., 2014; Zhang et al., 2019). Current findings were based on a task that used single-word stimuli and thus likely did not fully engage these areas. This consideration warrants a wider array of tasks to better characterize auditory processing within STGA, PP, and InsA. HGAL and STGM are the first auditory cortical areas to exhibit alpha suppression and thus are putative targets of modulatory influences from higher cortical areas. Future analysis of functional and effective connectivity using iEEG will provide needed insights into the origin of these modulatory top-down signals and further elucidate the dynamics of feedforward and feedback information flow within the human cortex (Banks et al., 2020). Ultimately, this line of research will refine the hierarchical organization of human auditory cortex.

References

Angulo-Perkins A, Aubé W, Peretz I, Barrios FA, Armony JL, Concha L (2014) Music listening engages specific cortical regions within the temporal lobes: differences between musicians and non-musicians. Cortex 59:126–137.

Bamiou DE, Musiek FE, Luxon LM (2003) The insula (Island of Reil) and its role in auditory processing. Literature review. Brain Res Brain Res Rev 42:143–154.

Banks MI, Krause BM, Endemann CM, Campbell DI, Kovach CK, Dyken ME, Kawasaki H, Nourski KV (2020) Cortical functional connectivity indexes arousal state during sleep and anesthesia. Neuroimage 211:116627.

Baş ar E, Güntekin B (2012) A short review of alpha activity in cognitive processes and in cognitive impairment. Int J Psychophysiol 86:25–38.

Bauer M, Stenner MP, Friston KJ, Dolan RJ (2014) Attentional modulation of alpha/beta and gamma oscillations reflect functionally distinct processes. J Neurosci 34:16117–16125.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical andpowerful approach to multiple testing. J Royal Stat Soc Series B Stat Methodol 57:289300.

Billig AJ, Herrmann B, Rhone AE, Gander PE, Nourski KV, Snoad BF, Kovach CK, Kawasaki H, Howard MA, 3rd, Johnsrude IS (2019) A Sound-sensitive source of alpha oscillations in human non-primary auditory cortex. J Neurosci 39:8679–8689.

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.

Blenkmann AO, Collavini S, Lubell J, Llorens A, Funderud I, Ivanovic J, Larsson PG, Meling TR, Bekinschtein T, Kochen S, Endestad T, Knight RT, Solbakk AK (2019) Auditory deviance detection in the human insula: an intracranial EEG study. Cortex 121:189–200.

Boemio A, Fromm S, Braun A, Poeppel D (2005) Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat Neurosci 8:389–395.

Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA 3rd (2009) Coding of repetitive transients by auditory cortex on Heschl's gyrus. J Neurophysiol 102:2358–2374.

Brunel N, Wang XJ (2003) What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. J Neurophysiol 90:415–430.

Burton H, Jones EG (1976) The posterior thalamic region and its cortical projection in New World and Old World monkeys. J Comp Neurol 168:249–301.

Camalier CR, D'Angelo WR, Sterbing-D'Angelo SJ, de la Mothe LA, Hackett TA (2012) Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. Proc Natl Acad Sci U S A 109:18168–18173.

Chang EF, Raygor KP, Berger MS (2015) Contemporary model of language organization: an overview for neurosurgeons. J Neurosurg 122:250–261.

Chao ZC, Takaura K, Wang L, Fujii N, Dehaene S (2018) Large-scale cortical networks for hierarchical prediction and prediction error in the primate brain. Neuron 100:1252–1266.e3.

Crone NE, Boatman D, Gordon B, Hao L (2001) Induced electrocorticographic gamma activity during auditory perception. Brazier Award-winning article, 2001. Clin Neurophysiol 112:565–582.

Crone NE, Korzeniewska A, Franaszczuk PJ (2011) Cortical γ responses: searching high and low. Int J Psychophysiol 79:9–15.

de Heer WA, Huth AG, Griffiths TL, Gallant JL, Theunissen FE (2017) The hierarchical cortical organization of human speech processing. J Neurosci 37:6539–6557.

Dimitrijevic A, Smith ML, Kadis DS, Moore DR (2017) Cortical alpha oscillations predict speech intelligibility. Front Hum Neurosci 11:88.

Forseth KJ, Hickok G, Rollo PS, Tandon N (2020) Language prediction mechanisms in human auditory cortex. Nat Commun 11:5240.

Friederici AD, Meyer M, von Cramon DY (2000) Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. Brain Lang 75:289–300.

Garofolo J, Lamel L, Fisher W, Fiscus J, Pallett D, Dahlgren N (1993) Darpa timit acoustic-phonetic continuous speech corpus CD-ROM {TIMIT}. Gaithersburg, MD: NISTIR 4930.

Geschwind N, Levitsky W (1968) Human brain: left-right asymmetries in temporal speech region. Science 161:186–187.

Goldman RI, Stern JM, Engel J Jr, Cohen MS (2002) Simultaneous EEG and fMRI of the alpha rhythm. Neuroreport 13:2487–2492.
Griffiths TD, Warren JD (2002) The planum temporale as a computational hub. Trends Neurosci 25:348–353.
Guo N, Si X, Zhang Y, Ding Y, Zhou W, Zhang D, Hong B (2021) Speech frequency-following response in human auditory cortex is more than a simple tracking. Neuroimage 226:117545.
Habib M, Daquin G, Milandre L, Royere ML, Rey M, Lanteri A, Salamon G, Khalil R (1995) Mutism and auditory agnosia due to bilateral insular damage–role of the insula in human communication. Neuropsychologia 33:327–339.
Hackett TA (2015) Anatomic organization of the auditory cortex. Handb Clin Neurol 129:27–53.
Häkkinen S, Rinne T (2018) Intrinsic, stimulus-driven and task-dependent connectivity in human auditory cortex. Brain Struct Funct 223:2113–2127.
Hamilton LS, Edwards E, Chang EF (2018) A spatial map of onset and sustained responses to speech in the human superior temporal gyrus. Curr Biol 28:1860–1871.e4.
Hamilton LS, Oganian Y, Hall J, Chang EF (2021) Parallel and distributed encoding of speech across human auditory cortex. Cell 184:4626–4639.e13.
Hickok G (2009) The functional neuroanatomy of language. Phys Life Rev 6:121–143.
Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci 8:393–402.
Hickok G, Poeppel D (2015) Neural basis of speech perception. Handb Clin Neurol 129:149–160.
Humphries C, Love T, Swinney D, Hickok G (2005) Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. Hum Brain Mapp 26:128–138.
Humphries C, Sabri M, Lewis K, Liebenthal E (2014) Hierarchical organization of speech perception in human auditory cortex. Front Neurosci 8:406.
Hyde KL, Peretz I, Zatorre RJ (2008) Evidence for the role of the right auditory cortex in fine pitch resolution. Neuropsychologia 46:632–639.
Jäncke L, Wüstenberg T, Scheich H, Heinze HJ (2002) Phonetic perception and the temporal cortex. Neuroimage 15:733–746.
Jenkinson M, Smith S (2001) A global optimisation method for robust affine registration of brain images. Med Image Anal 5:143–156.
Jensen O, Mazaheri A (2010) Shaping functional architecture by oscillatory alpha activity: gating by inhibition. Front Hum Neurosci 4:186.
Klein A, Tourville J (2012) 101 labeled brain images and a consistent human cortical labeling protocol. Front Neurosci 6:171.
Klimesch W, Sauseng P, Hanslmayr S (2007) EEG alpha oscillations: the inhibition-timing hypothesis. Brain Res Rev 53:63–88.
Kovach CK, Gander PE (2016) The demodulated band transform. J Neurosci Methods 261:135–154.
Leaver AM, Rauschecker JP (2010) Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. J Neurosci 30:7604–7612.
Leff AP, Schofield TM, Stephan KE, Crinion JT, Friston KJ, Price CJ (2008) The cortical dynamics of intelligible speech. J Neurosci 28:13209–13215.
Loftus WC, Tramo MJ, Thomas CE, Green RL, Nordgren RA, Gazzaniga MS (1993) Three-dimensional quantitative analysis of hemispheric asymmetry in the human superior temporal region. Cereb Cortex 3:348–355.
Mahjoory K, Schoffelen JM, Keitel A, Gross J (2020) The frequency gradient of human resting-state brain oscillations follows cortical hierarchies. Elife 9:e53715.
Mesgarani N, Chang EF (2012) Selective cortical representation of attended speaker in multi-talker speech perception. Nature 485:233–236.
Mesgarani N, Cheung C, Johnson K, Chang EF (2014) Phonetic feature encoding in human superior temporal gyrus. Science 343:1006–1010.
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.
Nagahama Y, Schmitt AJ, Nakagawa D, Vesole AS, Kamm J, Kovach CK, Hasan D, Granner M, Dlouhy BJ, Howard MA, Kawasaki H (2018a) Intracranial EEG for seizure focus localization: evolving techniques, outcomes, complications, and utility of combining surface and depth electrodes. J Neurosurg 1–13.
Nagahama Y, Schmitt AJ, Dlouhy BJ, Vesole AS, Gander PE, Kovach CK, Nakagawa D, Granner MA, Howard MA, Kawasaki H (2018b) Utility and safety of depth electrodes within the supratemporal plane for intracranial EEG. J Neurosurg 131:772–780.
Nourski KV (2017) Auditory processing in the human cortex: an intracranial electrophysiology perspective. Laryngoscope Investig Otolaryngol 2:147–156.
Nourski KV, Brugge JF (2011) Representation of temporal sound features in the human auditory cortex. Rev Neurosci 22:187–203.
Nourski KV, Brugge JF, Reale RA, Kovach CK, Oya H, Kawasaki H, Jenison RL, Howard MA 3rd (2013) Coding of repetitive transients by auditory cortex on posterolateral superior temporal gyrus in humans: an intracranial electrophysiology study. J Neurophysiol 109:1283–1295.
Nourski KV, Howard MA 3rd (2015) Invasive recordings in the human auditory cortex. Handb Clin Neurol 129:225–244.
Nourski KV, Reale RA, Oya H, Kawasaki H, Kovach CK, Chen H, Howard MA, 3rd, Brugge JF (2009) Temporal envelope of time-compressed speech represented in the human auditory cortex. J Neurosci 29:15564–15574.
Nourski KV, Steinschneider M, McMurray B, Kovach CK, Oya H, Kawasaki H, Howard MA 3rd (2014) Functional organization of human auditory cortex: investigation of response latencies through direct recordings. Neuroimage 101:598–609.
Nourski KV, Steinschneider M, Rhone AE, Oya H, Kawasaki H, Howard MA 3rd., McMurray B (2015) Sound identification in human auditory cortex: differential contribution of local field potentials and high gamma power as revealed by direct intracranial recordings. Brain Lang 148:37–50.
Nourski KV, Steinschneider M, Rhone AE (2016) Electrocorticographic activation within human auditory cortex during dialog-based language and cognitive testing. Front Hum Neurosci 10:202.
Nourski KV, Steinschneider M, Rhone AE, Howard MA III (2017) Intracranial electrophysiology of auditory selective attention associated with speech classification tasks. Front Hum Neurosci 10:691.
Nourski KV, Steinschneider M, Rhone AE, Kovach CK, Banks MI, Krause BM, Kawasaki H, Howard MA (2021) Electrophysiology of the human superior temporal sulcus during speech processing. Cereb Cortex 31:1131–1148.
Patel P, Long LK, Herrero JL, Mehta AD, Mesgarani N (2018) Joint representation of spatial and phonetic features in the human core auditory cortex. Cell Rep 24:2051–2062.e2.
Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci 12:718–724.
Scott SK, McGettigan C (2013) Do temporal processes underlie left hemisphere dominance in speech perception? Brain Lang 127:36–45.
Shahin AJ, Alain C, Picton TW (2006) Scalp topography and intracerebral sources for ERPs recorded during auditory target detection. Brain Topogr 19:89–105.
Steinmetz H, Rademacher J, Jäncke L, Huang YX, Thron A, Zilles K (1990) Total surface of temporoparietal intrasylvian cortex: diverging left-right asymmetries. Brain Lang 39:357–372.
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 3rd (2011) Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. Cereb Cortex 21:2332–2347.
Steinschneider M, Nourski KV, Fishman YI (2013) Representation of speech in human auditory cortex: is it special? Hear Res 305:57–73.

Steinschneider M, Nourski KV, Rhone AE, Kawasaki H, Oya H, Howard MA 3rd (2014) Differential activation of human core, non-core and auditory-related cortex during speech categorization tasks as revealed by intracranial recordings. Front Neurosci 8:240.

Traub RD, Cunningham MO, Gloveli T, LeBeau FE, Bibbig A, Buhl EH, Whittington MA (2003) GABA-enhanced collective behavior in neuronal axons underlies persistent gamma-frequency oscillations. Proc Natl Acad Sci U S A 100:11047–11052.

Upadhyay J, Silver A, Knaus TA, Lindgren KA, Ducros M, Kim DS, Tager-Flusberg H (2008) Effective and structural connectivity in the human auditory cortex. J Neurosci 28:3341–3349.

Vouloumanos A, Kiehl KA, Werker JF, Liddle PF (2001) Detection of sounds in the auditory stream: event-related fMRI evidence for differential activation to speech and nonspeech. J Cogn Neurosci 13:994–1005.

Woolnough O, Forseth KJ, Rollo PS, Tandon N (2019) Uncovering the functional anatomy of the human insula during speech. Elife 8: e53086.

Yuan H, Liu T, Szarkowski R, Rios C, Ashe J, He B (2010) Negative covariation between task-related responses in alpha/beta-band activity and BOLD in human sensorimotor cortex: an EEG and fMRI study of motor imagery and movements. Neuroimage 49:2596–2606.

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.

Zhang Y, Zhou W, Wang S, Zhou Q, Wang H, Zhang B, Huang J, Hong B, Wang X (2019) The roles of subdivisions of human insula in emotion perception and auditory processing. Cereb Cortex 29:517–528.