NeuroImage 202 (2019) 116076

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Direct electrophysiological mapping of human pitch-related processing in auditory cortex

Phillip E. Gander [a,*], Sukhbinder Kumar [b,c], William Sedley [b], Kirill V. Nourski [a,d], Hiroyuki Oya [a], Christopher K. Kovach [a], Hiroto Kawasaki [a], Yukiko Kikuchi [b], Roy D. Patterson [e], Matthew A. Howard III [a,d,f], Timothy D. Griffiths [a,b,c]

[a] Human Brain Research Laboratory, Department of Neurosurgery, The University of Iowa, Iowa City, IA, USA
[b] Institute of Neuroscience, Newcastle University Medical School, Newcastle, UK
[c] Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, UK
[d] Iowa Neuroscience Institute, The University of Iowa, Iowa City, IA, USA
[e] Department of Physiology, Development and Neuroscience, University of Cambridge, UK
[f] Pappajohn Biomedical Institute, The University of Iowa, Iowa City, IA, USA

ARTICLE INFO

Keywords:
Electrocorticography
High gamma
Neurophysiology
Evoked
Induced

ABSTRACT

This work sought correlates of pitch perception, defined by neural activity above the lower limit of pitch (LLP), in auditory cortical neural ensembles, and examined their topographical distribution. Local field potentials (LFPs) were recorded in eight patients undergoing invasive recordings for pharmaco-resistant epilepsy. Stimuli consisted of bursts of broadband noise followed by regular interval noise (RIN). RIN was presented at rates below and above the LLP to distinguish responses related to the regularity of the stimulus and the presence of pitch itself. LFPs were recorded from human cortical homologues of auditory core, belt, and parabelt regions using multicontact depth electrodes implanted in Heschl's gyrus (HG) and Planum Temporale (PT), and subdural grid electrodes implanted over lateral superior temporal gyrus (STG). Evoked responses corresponding to the temporal regularity of the stimulus were assessed using autocorrelation of the evoked responses, and occurred for stimuli below and above the LLP. Induced responses throughout the high gamma range (60–200 Hz) were present for pitch values above the LLP, with onset latencies of approximately 70 ms. Mapping of the induced responses onto a common brain space demonstrated variability in the topographical distribution of high gamma responses across subjects. Induced responses were present throughout the length of HG and on PT, which is consistent with previous functional neuroimaging studies. Moreover, in each subject, a region within lateral STG showed robust induced responses at pitch-evoking stimulus rates. This work suggests a distributed representation of pitch processing in neural ensembles in human homologues of core and non-core auditory cortex.

1. Introduction

Pitch analysis is a fundamental aspect of human audition; it signals prosody information in speech, melody in music, and source location in auditory scene analysis. Considerable advances in our understanding of where pitch processing takes place in the auditory pathway have been made. Responses to missing fundamental harmonic complex tones, consistent with a spatiotemporal code for pitch, have been found at the level of the auditory nerve (Cedolin and Delgutte, 2010). In the cortex, results from multiple recent studies support the existence of a pitch-sensitive non-primary auditory cortical region. However, debate still exists regarding the organization and location of specialized cortical neurons devoted to the processing of pitch. (For reviews of animal neurophysiology and human imaging studies, see Wang and Walker, 2012, and Griffiths and Hall, 2012, respectively). Using single-cell recordings in marmoset auditory cortex, Bendor and Wang (2005)

* Corresponding author. Department of Neurosurgery, University of Iowa Hospitals and Clinics, 200 Hawkins Dr., Iowa City, IA, 52242, USA.
E-mail addresses: phillip-gander@uiowa.edu (P.E. Gander), sukhbinder.kumar@newcastle.ac.uk (S. Kumar), william.sedley@newcastle.ac.uk (W. Sedley), kirill-nourski@uiowa.edu (K.V. Nourski), hiroyuki-oya@uiowa.edu (H. Oya), christopher-kovach@uiowa.edu (C.K. Kovach), hiroto-kawasaki@uiowa.edu (H. Kawasaki), yukiko-kikuchi@newcastle.ac.uk (Y. Kikuchi), rdp1@cam.ac.uk (R.D. Patterson), matthew-howard@uioaw.edu (M.A. Howard), tim.griffiths@newcastle.ac.uk (T.D. Griffiths).

https://doi.org/10.1016/j.neuroimage.2019.116076
Received 5 September 2018; Received in revised form 28 July 2019; Accepted 5 August 2019
Available online 8 August 2019

observed pitch-selective responses to missing fundamental harmonic complex sounds in a low-frequency tuned region that overlapped a border between primary and non-primary auditory cortex. Whether pitch processing requires a single specialized cortical region or distributes coding over multiple neural regions is the topic of ongoing debate. Recordings from ferrets performing a pitch discrimination task found neurons distributed throughout auditory cortex that represented pitch-related fundamental frequency information (Bizley et al., 2010, 2013). By manipulating periodicity through changes in repetition rate of pitch-evoking pulse train stimuli, broadly distributed responses have also been observed in auditory cortex of marmoset monkeys (Bendor and Wang, 2010). Importantly however, a subset of these neurons, located in a region similar to that observed in Bendor and Wang (2005), were found to also be responsive to a manipulation of pitch salience through changes in temporal regularity (jittered or not). Two fundamental aspects of the neural coding of pitch remain uncertain: the character of the neural responses that underlie it, and whether these responses include activity within a single 'pitch center' or more widely distributed as part of a 'pitch system'. For auditory neuroscience to determine how pitch perception occurs, part of the answer likely depends on whether or how it is a system that is dependent on a number of aspects that are interdependent, much like how color vision is an interconnected system and is not just based on processing in V4 (Conway, 2009).

The use of animal models of pitch allows detailed understanding afforded by systematic invasive recordings from multiple locations, which establishes models that can be tested in humans. While most studies of pitch processing in humans rely on non-invasive brain imaging, a handful have taken advantage of opportunities to record directly from the brain in neurosurgical patients for whom such recording are clinically justified. Early studies using both non-invasive and invasive techniques in humans identified a response to the onset of a pitch stimulus referred to as regular interval noise (RIN) that is different from the response to the onset of sound (Krumbholz et al., 2003; Schonwiesner and Zatorre, 2008), which localized to medial Heschl's gyrus (HG) (Krumbholz et al., 2003) and planum temporale (PT) adjacent to lateral HG (Schonwiesner and Zatorre, 2008), corresponding to primary and non-primary auditory cortex, respectively (e.g., (Galaburda and Sanides, 1980; Hackett et al., 2001; Liegeois-Chauvel et al., 1991; Morosan et al., 2001; Talavage et al., 2000). In a subsequent study of two subjects, Griffiths et al. (2010) recorded LFPs from auditory cortex in response to stimuli above and below the lower limit of pitch (LLP), which was defined as the lowest frequency where you can reliably detect semi-tone changes in the notes of a short melody. The LLP is close to 30 Hz in both humans and macaques (Krumbholz et al., 2000; Pressnitzer et al., 2001; Joly et al., 2014). A dissociation was demonstrated between responses that were phase-locked to stimulus regularity that occurs both above and below the LLP and non-phase-locked high gamma responses (80–120 Hz) that occurred only when the repetition rate was above the LLP. The latter response is, accordingly, a candidate neural correlate of pitch processing inasmuch as it only occurs when there is a concomitant perception of pitch. The non-phase-locked response localized to both medial and lateral portions of HG, thus spanning core as well as non-core auditory cortex.

Human invasive recordings (e.g., Griffiths et al., 2010; Schonwiesner and Zatorre, 2008) are necessarily confined to regions in which recordings have a clinical rationale and are otherwise limited by constraints of sampling and accessibility (Nagahama et al., 2018). Coverage of auditory cortices in these previous studies amounts to 32 recording locations over three subjects, providing little basis for insight about spatial distribution. In contrast, non-invasive functional MRI (fMRI) imaging of the blood oxygenation level-dependent (BOLD) response, allows whole brain mapping including the auditory cortex in the superior temporal plane (STP). Imaging studies have used a variety of pitch-associated stimuli and criteria for pitch responses. Early studies demonstrated maximal responses to pitch-associated stimuli in lateral HG (Patterson et al., 2002); although there were also significant responses in medial HG (see Griffiths et al., 2010, for a discussion). Further studies have also implicated anterolateral PT, posterior to HG (e.g., Plack et al., 2014) and areas anterior to HG (Barrett and Hall, 2006; Norman-Haignere et al., 2013; Puschmann et al., 2010).

The present study uses a perceptual criterion for pitch processing based on psychophysics in humans, i.e., human observers consider a sound to have a pitch-like quality when it has a repetition rate above approximately 30 Hz. The repetition rate gives the sound its 'pitch value', rather than a sound quality that is considered to be 'beating' when it is below 30 Hz. By combining the high spatiotemporal resolution and signal to noise ratio of direct recordings with a high spatial resolution allows comparison with fMRI results. The study tests the relationship between LFP high gamma band responses and pitch processing in data recorded from 122 depth electrode contacts in the superior temporal plane over eight patients, along with recordings from 622 electrode contacts in grid arrays over the lateral convexity of the temporal lobe. This work provides a detailed test of the hypothesis that gamma band responses that selectively occur for repetition rate stimuli above the LLP are neural correlates of pitch processing; in addition, it provides a rigorous test of the hypothesis that such responses are restricted to a small region within human auditory core and belt homologues. The results show that high gamma responses were restricted to stimulus rates above the LLP. The responses were observed within multiple regions of auditory cortex supporting a distribution of neural ensembles for pitch-related processing in human homologues of core and non-core auditory cortex, including locations on lateral STG not previously reported in human LFP work.

2. Materials and methods

2.1. Subjects

Subjects were eight neurosurgical patients (4 female, aged 28–56, median age 42) undergoing chronic invasive monitoring for medically refractory epilepsy to identify seizure foci prior to resection surgery. All research protocols were approved by the National Institutes of Health and The University of Iowa Institutional Review Board. Subjects provided written, informed consent to the research prior to surgery and could rescind consent at any time without detriment to their medical evaluation. Audiometric testing prior to surgery found the presence of either normal hearing or a mild loss (four subjects had 25–40 dB HL thresholds at three or fewer test frequencies at standard octave steps between 0.25 and 8 kHz). All subjects had Pure Tone Average hearing scores in the normal clinical range (<25 dB HL, average of 0.5, 1, 2, 4 kHz) and suprathreshold speech reception thresholds were greater than 88% in all ears; none used hearing assistance devices or commented on noticeable hearing impairment. Subject R180 was left-handed and the other subjects were right-hand dominant. Presurgical Wada test results confirmed left-hemisphere language dominance in all subjects except L162 with bilateral language. Data from two of the eight subjects were reported previously and are reanalyzed here (R154 and L156, Griffiths et al., 2010).

Electrodes were implanted only on the basis of clinical considerations, and in each subject consisted of coverage over portions of frontal, parietal, and temporal cortex. Three subjects were implanted on the right side (as indicated by the letter prefix 'R' in the subject code) and five on the left (prefix code 'L'). Only responses from auditory cortex (HG, PT, and STG) are reported here, and all data containing any potential epileptiform activity or interictal spikes were removed. All recordings were free of seizure activity at least 1 h before and after the experiment. All reported electrode contacts were confirmed not to be involved in the generation of epileptiform activity and did not overlie resected tissue.

2.2. Recording

Recordings were made under two slightly different experimental setups: one in the first group of subjects R154, L156, L162, L178, R180, R186, and another for L403 and L404; the differences are described

below. In all subjects, intracranial data were digitally recorded from depth electrodes (Ad-Tech Medical Instrument, Racine, WI) implanted along the long axis of HG (Howard et al., 1996; Nagahama et al., 2018; Reddy et al., 2010). In the first subject group, these electrodes included six low impedance clinical and 14–15 high impedance (70 - 300 kΩ) research contacts spaced at 2 mm intervals, giving an equal sampling of recording contacts across 34–36 mm (see Supplementary Table 1 for contact impedance type across all HG electrodes). The two most lateral low impedance contacts were used for reference and ground. Reconstructions of the HG contact locations are presented for each subject in Supplementary Fig. 1 (top row). The data from this group were acquired using a Tucker Davis Technologies (Alachua, FL) RZ2 system. In subjects L403 and L404 HG electrodes consisted of 12 low impedance clinical contacts spaced at 5 mm intervals. An additional electrode (8 contacts, 5 mm spacing) in L404 was placed along the STP orthogonal to the HG electrode and provided coverage of HG and PT. The electrode recordings in L403 and L404 were referenced to a subgaleal electrode and acquired using a Neuralynx (Bozeman, MT) Atlas system. In all subjects, recordings were simultaneously made from multicontact subdural grid electrodes (Ad-Tech Medical Instrument, Racine, WI) implanted over perisylvian cortex covering the posterior STG, which included the lateral aspect of HG. This lateral temporal grid electrode contained 96 platinum-iridium disc contacts (2.3 mm exposed diameter, 5 mm center-to-center distance) embedded in a silicon membrane in an 8 x 12 configuration (3.5 × 5.5 cm), referenced to a subgaleal electrode. For subjects L403 and L404 the silicon membrane was cut and slightly separated to allow for the placement of another depth electrode. In the first subject group data were acquired using a 1–1000 Hz bandpass filter (sampled at 2034 Hz low impedance and 12207 Hz high impedance) and 16-bit resolution with the RZ2 system and in subjects L403 and L404 data were acquired 0.1–500 Hz bandpass (sampled at 2000 Hz) at 24-bit resolution with the Atlas system. The electrodes remained in place between 2 and 4 weeks under clinical direction. Experiments took place in a patient suite, which had specialized electrical shielding.

2.3. Imaging

A high-resolution T1-weighted structural MRI of the brain was acquired for each subject before and after electrode implantation. In the first 6-subject group images were acquired from a 3T Siemens TIM Trio scanner with a 12-channel head coil (MPRAGE: 0.78 × 0.78 mm, slice thickness 1.0 mm, TR = 2.530 s, TE = 3.520 ms, average of two). In the second group of two subjects images were acquired from a 3T GE Discovery MR750w scanner with a 32-channel head coil (IR-FSPGR: 1 × 1 mm, slice thickness 0.8 mm, TR = 8.504 ms, TE = 3.288 ms). To determine the location of lateral temporal surface recording electrode contacts on the preoperative structural MRI, these images were coregistered to post-implantation structural MRIs using a 3D linear registration algorithm (Functional MRI of the Brain Linear Image Registration Tool, Jenkinson et al., 2012) and custom Matlab v.9.0 (MathWorks, Natick, MA) scripts using guidance from computed tomography (CT) scans (in-plane resolution 0.51 × 0.51 mm, slice thickness 1.0 mm). When possible, results were also compared with intraoperative photographs to ensure reconstruction accuracy. Each individual's brain was coregistered to the ICBM152 template (Fonov et al., 2011).

To facilitate the comparison of responses across different subjects while retaining the relative location of depth electrode contacts on individual STP anatomy, custom Matlab scripts were written to implement the following procedure: (1) manual selection of 50–100 control point pairs on the pre- and post-implantation images, (2) thin-plate spline warping of post-to pre-implantation image based on the control points, (3) coregistration of pre-implantation image to the CIT168 brain (Tyszka and Pauli, 2016) using ANTs software (Avants et al., 2007), (4) plotting of responses at each electrode contact at a 2 mm diameter. MNI coordinates for each STP contact location were determined from this procedure (see Supplementary Table 1). Based on individual anatomy, recording sites were only included for analysis if they were within or immediately adjacent STP gray matter; this removed 19/142 contacts (not shown). For display purposes the locations of the STP contacts are shown on the cortical surface, which modifies the localized z-plane coordinate.

The accuracy of this method was confirmed against a manual method in which each subject's STP was mapped onto the ICBM 152 template using the following procedure (Oya et al., 2014): (1) Surface reconstruction of whole hemispheres with FreeSurfer (Reuter et al., 2010); (2) STP patch creation; (3) Warping of individual subjects' STP patches onto the template brain using spherical diffeomorphic deformation (Yeo et al., 2010); (4) Projection of responses at each contact to the surface vertices. Notably, these electrode projection procedures involved warping of images and electrode contact locations onto the preoperative brain which introduced some small amount of error in localization. However, while the results between both methods were very similar (mean and standard deviation absolute difference in mm, X: 1.29 ± 1.23, Y: 2.51 ± 1.57, Z: 2.01 ± 1.45), we used the volumetric coregistration to the CIT168 atlas because it retained slightly more accurate projections of intercontact distances and trajectories of the depth electrode.

For each subject the anterolateral extent of core (primary) auditory cortex was approximated using physiological criteria based on presence of phase-locked responses to 100 Hz click trains and short-latency (<20 ms) peaks in the averaged auditory evoked potential (AEP) waveforms to these stimuli (Brugge et al., 2009; Nourski et al., 2016). Delineation of core and non-core auditory cortex along the HG electrode array in each subject projected onto the ICBM 152 template is presented in Supplementary Fig. 1.

2.4. Stimuli

Stimuli were sound complexes consisting of 1 s bursts of broadband noise followed by 1.5 s regular interval noise (Fig. 1 and Supplementary Fig. 2; see also Griffiths et al., 2010, for a detailed description of the stimuli). In brief, each of the six RIN stimuli were generated from a unique random noise exemplar using a delay-and-add algorithm with 16 iterations, highpass-filtered at 800 Hz. To mask cochlear distortion products, a broadband noise, normalized to the peak power spectral density of the filtered noise, was added below this cutoff. RIN segments were constructed with six repetition rates (8, 16, 32, 64, 128, 256 Hz), corresponding to values below (8 and 16 Hz), at the boundary of (32 Hz), and above (64, 128, 256 Hz) the LLP (Krumbholz et al., 2000; Pressnitzer

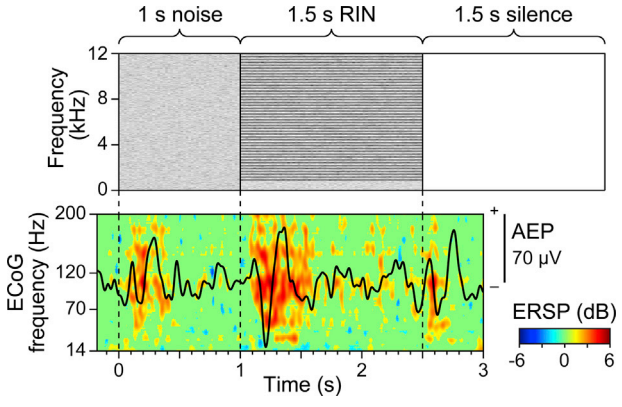

Fig. 1. Trial schematic and analysis example. A spectrogram is shown for a single stimulus trial and the resulting response representation for a single recording site. 1 s of broadband white noise is followed by 1.5 s of RIN at one of 6 repetition rates (8, 16, 32, 64, 128, or 256 Hz), followed by 1.5 s silence. The response to the most salient pitch (256 Hz RIN) is shown as the average auditory evoked potential (AEP, black line) and the induced Event Related Spectral Perturbation (ERSP) response shown in the time-frequency color representation using the scale on right.

et al., 2001). We used 16 iterations of the delay-and-add cycle in order to create a quasi-regular stimulus and a strong pitch percept.

Each sound was followed by 1.5 s of silence, with an additional random timing jitter between ±23 ms before the presentation of the next stimulus. Fifty repetitions of each sound complex were presented in random order.

2.5. Procedure

Recordings were made in a dedicated recording facility incorporating an electromagnetically shielded room. Subjects were awake and reclined in a hospital bed or an armchair and were instructed to listen to the sounds. The stimuli were delivered diotically via Etymotic ER4B earphones in custom earmolds at a comfortable level, adjusted for each subject to be in the range of 45–55 dB above threshold.

2.6. Analysis

Analyses were performed using EEGLAB (Delorme and Makeig, 2004) and custom Matlab scripts for statistical testing. Analysis of LFP responses included assessment of induced, non-phase-locked event-related spectral perturbation (ERSP) using wavelet spectral decomposition and evoked, phase-locked AEPs. All data were downsampled to 1000 Hz and any line noise was removed using custom Matlab scripts implementing the demodulated band transform (Kovach and Gander, 2016). Trials were epoched at −300 to 3400 ms with respect to sound complex onset. Artifact rejection was first performed by manual inspection of continuous ECoG to remove any trials that contained low amplitude spiking artifacts, and then run through automatic rejection for amplitude and kurtosis. The average number of retained trials was 261 (220–299) of the total of 300.

For time-domain displayed AEP analyses, raw data were band-pass filtered in the range of 1–20 Hz. For each trial, mean voltage within a −200 to 0 ms pre-stimulus baseline was subtracted before averaging across trials. Root mean square power of the evoked response was calculated to provide a measure of the energy during the steady state of the noise segment (800–1000 ms), and after the transition to the RIN segment (1050–1250 ms) which for pitch-evoking stimuli has been termed the pitch onset response (Schonwiesner and Zatorre, 2008; Krumbholz et al., 2003). To detect activity related to the temporal regularity of the stimulus, autocorrelation analyses were performed by taking unfiltered AEP data that were then highpass filtered at 0.5 Hz to remove the influence of the high autocorrelation at zero delay for the 256 Hz RIN stimulus delay value at 4 ms. The data were then 1/f whitened (by taking the first order differential) for each repetition rate. The maximum value was taken within a window ± 2 ms from the delay corresponding to each stimulus repetition rate, except for the 256 Hz condition, which was taken at 4 ms to avoid potentially inflated values at 2 or 3 ms. Significant responses for each subject were determined from a permutation test (1000 shuffles) of the autocorrelation data. For each shuffle, each trial was randomly shifted + or – 125 ms (an integer multiple of the stimulus period for each rate) thereby removing the inherent phase locking while retaining the overall structure of the data. The significance threshold for each subject was taken as the maximum value at any electrode contact and condition from this procedure. To display the degree of response phase-locking across trials, inter-trial (phase) coherence (ITC) was calculated on the unfiltered AEP between 14 and 200 Hz (see Supplementary Fig. 3), applying the EEGLAB permutation test (shuffling timepoints and trials) with a significance threshold of 0.01 (two-tailed); all nonsignificant values are colored green. This permutation test is uncorrected (i.e., mean data value) and for display purposes only.

Induced activity was analyzed using Morlet wavelets based on a −300 to −100 ms pre-stimulus reference period and a sliding wavelet window size with linearly increasing cycles across 14–200 Hz (3–21.4 cycles) for ERSP calculation. The lower bound of 14 Hz was determined from the lowest end of a standard frequency band at which the number of cycles in the available baseline period was sufficient to estimate the power using a wavelet analysis. An example averaged evoked and induced response for the 256 Hz RIN condition from one contact is shown in Fig. 1. In the time-frequency plot of Fig. 1 and Supplementary Figs. 1, 3, and 5, the EEGLAB permutation test (for display purposes) was applied using a significance threshold of 0.01 (two-tailed); all nonsignificant values are colored green. This (uncorrected) permutation test shuffles timepoints and trials across each frequency using the bandwidth of the wavelet applied in the time-frequency decomposition, comparing the response for the whole trial to the pre-stimulus reference period (−300 to −100 ms). To detect changes following the transition from noise to RIN the power in the 70–120 Hz gamma band was calculated for three consecutive 500 ms segments after the transition (0–500, 500–1000, 1000–1500), each referenced to the power during the 500 ms preceding the transition. These data were then subjected to a permutation analysis (1000 shuffles, corrected against the full data set) by randomizing the time windows of analysis between the 500 ms preceding the transition and the respective 500 ms segment following the transition (0–500, 500–1000, 1000–1500). Significance was determined at $p < 0.05$ (two-tailed). The more stringent, corrected permutation test for the induced gamma band activity was performed for hypothesis testing. High gamma responses recorded from the lateral temporal grid electrode were visualized on each subject's lateral surface reconstruction by smoothing of responses at each contact using a linear interpolation across grid contact locations.

To address potential activations in relation to salient slow spectro-temporal modulations that arise from RIN stimuli constructed with greater regularity (16 iterations, Barker et al., 2012), we performed an additional analysis using a subtraction of responses to the 16 Hz RIN from responses to 32, 64, 128, and 256 Hz RIN. This stimulus contains slow spectro-temporal fluctuations that are not present in the noise stimulus, and hence controls for the suggestion that responses to RIN might be associated with such stimulus changes or their perceptual timbral correlate. The analysis was otherwise identical to that described for induced activity above, except that permutation analyses used the same time windows rather than a fixed reference to the end of the noise period (see Supplementary Table 2).

3. Results

3.1. Evoked and induced responses at individual recording sites in STP

The panels of Fig. 2 present the data as repetition-rate histograms for each recording site within the STP in each subject. Evoked and induced responses are plotted in panels on the left and right, respectively, and individual subjects are presented in separate rows. Recordings were possible from 122 contacts. Supplementary Fig. 1 shows the detailed evoked and induced responses for one repetition rate below the LLP (16 Hz) and one above the LLP (256 Hz).

The evoked activity, calculated as the autocorrelation of the AEP waveform (Fig. 2: left column) at each RIN repetition rate shows a response pattern across subjects with peak activity at 32 or 64 Hz, less activity at 8, 16, and 128 Hz, and no response at 256 Hz. This pattern agrees with the 40 Hz maximum observed for auditory envelope following responses (Galambos et al., 1981; Ross et al., 2000) and the limits of the auditory frequency following response in HG (Brugge et al., 2009; Nourski et al., 2013). The majority of the subjects (6/8) showed evoked responses medial to the HG-core boundary as opposed to more lateral, which is compatible with the stronger stimulus locking known to exist in core auditory cortex (Brugge et al., 2009).

The induced activity patterns (Fig. 2: right column) demonstrate a significant response above the LLP in 79/122 contacts as compared to 8/122 below, during the first 500 ms after RIN onset. The total number of significant electrode contacts for each condition during the first 500 ms of RIN were as follows: 8 Hz = 1, 16 Hz = 8, 32 Hz = 18, 64 Hz = 53, 128 Hz = 71, 256 Hz = 75 (see Supplementary Table 2 for later time windows). Significant responses were distributed across HG showing no

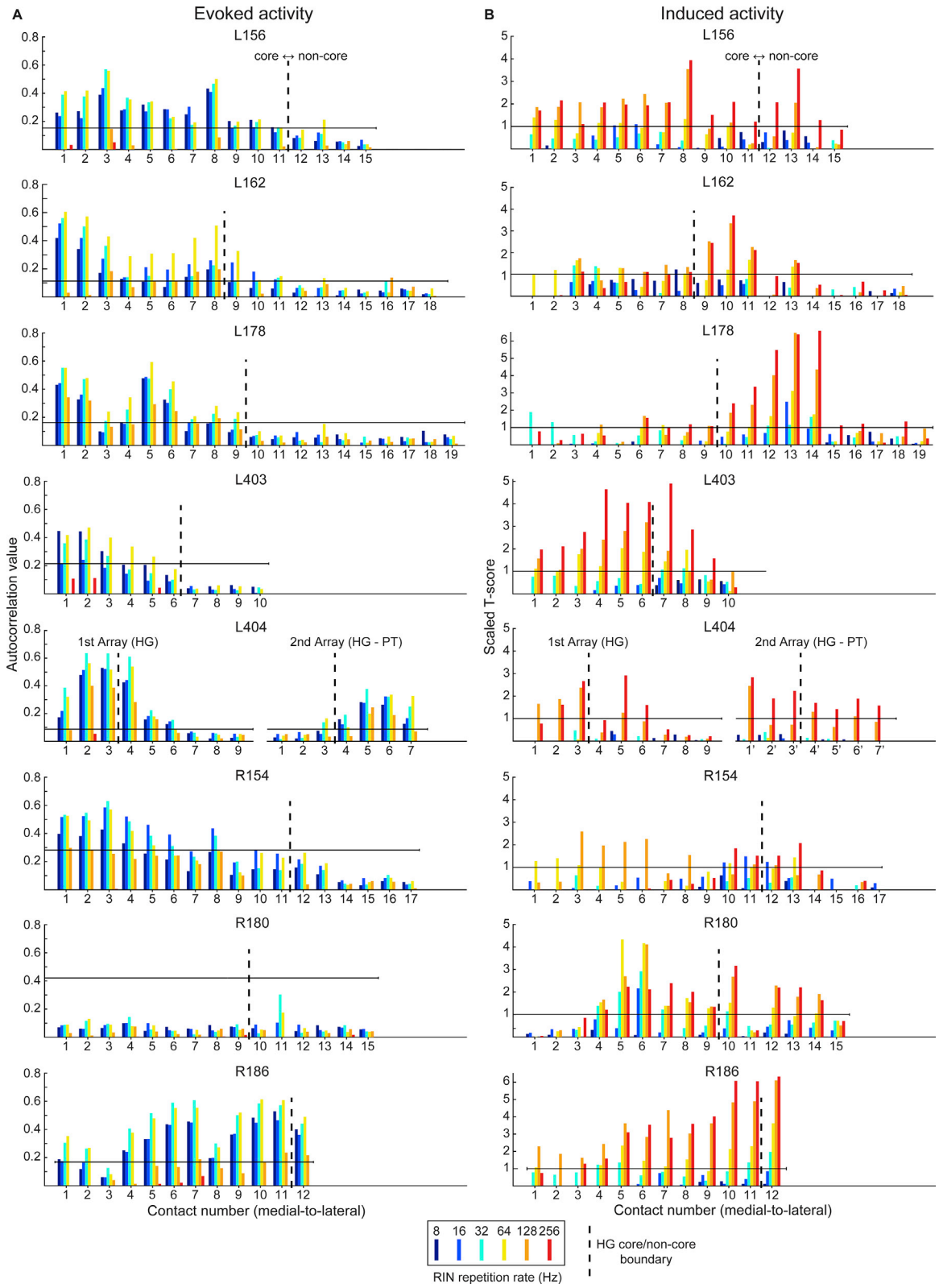

Fig. 2. Heschl's gyrus depth electrodes: details of evoked and induced activity. Evoked and induced responses are shown for each subject for each RIN condition for all depth electrode contacts (see Supplementary Table 1). Evoked activity is taken from the peak of the response autocorrelation averaged across all electrodes for each RIN repetition rate. Horizontal lines represent significance thresholds determined from permutation analysis for each subject. Induced activity is the ERSP high gamma (70–120 Hz) power from 0 to 500 ms after transition to the RIN stimulus referenced to the 500 ms preceding the transition. Horizontal lines represent significance thresholds determined from a permutation of *t*-scores scaled to the maximum for each subject. A dashed line represents the physiologically defined boundary of core auditory cortex based on induced and frequency following response to 100 Hz click train stimuli recorded in the same subject. Responses to the left/medial of the boundary are in auditory core and to the right/lateral are in non-core.

clear pattern with respect to the core/non-core boundary.

Supplementary Fig. 1 shows that along HG we found robust low-frequency evoked and high-frequency induced responses to the strongest pitch evoking 256 Hz RIN following the transition from noise in all eight subjects. This is in marked contrast to the largely non-existent evoked and induced responses to the transition from noise to 16 Hz RIN (see Supplementary Table 2). Onset latencies for auditory evoked responses to the RIN stimuli were only consistently measurable in the 128 and 256 Hz conditions peaking approximately 150–200 ms, with a shortening of latency for the 256 Hz stimulus (see Krumbholz et al., 2003). Induced response latency to the RIN stimulus was consistent across conditions and typically occurred around 70 ms (for those responses that were present, which typically emerged at 32 or 64 Hz). Induced responses to the 256 Hz RIN were observed in the high gamma range (60–200 Hz), however the magnitude and bandwidth of the response differed across subjects, as did the number and location of the electrode contacts in which this response was present. In comparison to the induced response to the noise, the high gamma response to RIN was more sustained, and in a greater number of contacts occurred for the duration of the stimulus (see Supplementary Table 2). This is reflected in significant responses across 101 contacts from 500 to 1000 ms and 132 contacts from 1000 to 1500 ms following the transition to RIN, compared to 24 significant responses to the noise stimulus (comparing the last 500 ms of noise to the 500–1000 ms window of the RIN stimulus). Onset and offset for both evoked and induced responses to the sound complex (1 s noise plus 1.5 s RIN) were similar in all RIN stimulus conditions.

Evoked activity was also assessed using inter-trial coherence (ITC). Supplementary Fig. 3 shows ITC for each subject for all depth electrode contacts for the 16 and 256 Hz RIN conditions. Unlike the autocorrelation analysis, this analysis demonstrates time locked activity at all frequencies (14–200 Hz) and not just the repetition rate of the RIN. The results show clear phase-locking to the stimulus complex onset across many HG contacts in most subjects independent of the stimulus condition. When present, the ITC response occurred throughout the stimulus, broadly across the displayed frequency range from approximately 14-200 Hz. During the RIN segment, weak phase-locking in the frequency range (typically a band of ±10 Hz) of the repetition rate was observed in some contacts for the 32, 64, and 128 Hz conditions (not shown) and some of their harmonics. For the 8 and 16 Hz stimuli, some contacts had a response in the upper harmonics, as can be observed for subject R186 in the 16 Hz ITC responses around 48, 64, and 80 Hz. No clear pattern was observed with respect to the transition to RIN at 1 s.

3.2. Evoked and induced responses in STP as a function of RIN repetition rate

Fig. 3A and Fig. 3B summarize evoked and induced responses, respectively, from all contacts in HG in each subject, separated by location with respect to the core (posteromedial)/non-core (anterolateral) boundary along HG. Specifically, Fig. 3A shows autocorrelation magnitude as a function of repetition rate, across all the HG electrode contacts. Significant responses occur both below and above the LLP. Note the significance thresholds are different than in Fig. 2; the responses at each contact are scaled for each subject to the value obtained from the permutation procedure which took the highest value for any condition in any contact. Supplementary Fig. 4 shows the pattern of evoked AEP pitch onset responses across STP contacts for the different stimulus conditions, which demonstrates a take-off above the LLP in each subject with maximum responses for the 256 Hz RIN stimulus. Fig. 3B shows the induced gamma activity in the same format as Fig. 3A. A similar pattern of activity is seen in posteromedial and anterolateral HG; mean activity is significant for the 128 and 256 Hz RIN conditions. Average activity across electrode contacts within each subject (not shown) demonstrated no significant responses at 8, 16, or 32 Hz in any subject, however seven of the eight subjects exhibited significant gamma band responses above the LLP; the totals were 64 Hz = 3, 128 Hz = 6, 256 Hz = 7.

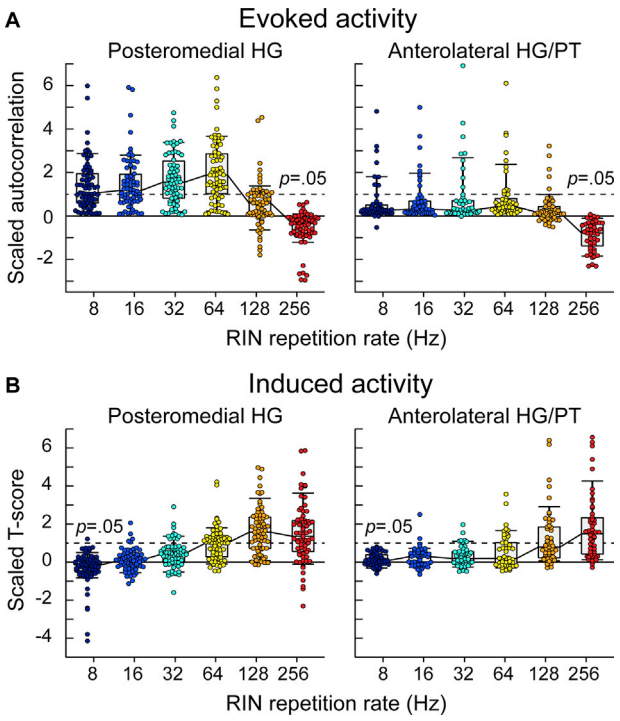

Fig. 3. Summary of evoked and induced activity within STP. Data are summarized as box-and whiskers plots across RIN repetition rates for all 122 STP contacts in the eight subjects. Each plot shows the median and the 10th, 25th, 75th, 90th percentiles; a line connects the median values across the six conditions. In separate graphs, the data are shown according to the contact location either posteromedial (core, n = 71) or anterolateral/PT (non-core, n = 51) to the physiologically defined boundary in each subject. A. Evoked activity is shown for the peak of the autocorrelation function of the response at each contact. Data for each subject are scaled to their maximum autocorrelation value from the shuffled permutation procedure. The dashed horizontal lines (at 1) represent the significance threshold. B. Induced activity is shown for the ERSP high gamma (70–120 Hz) power from 0 to 500 ms after transition to the RIN stimulus referenced to the 500 ms preceding the transition at each contact. Dashed horizontal lines (at 1) represent the significance threshold determined from a permutation of *t*-scores scaled to the maximum for each subject.

3.3. Control for timbral properties of RIN

To address the possibility that these responses might be due to slow spectro-temporal modulations (Hall and Plack, 2009), we subtracted the response to the 16 Hz RIN from the response at each higher rate; these results are plotted in Supplementary Fig. 5. Because the slow spectro-temporal modulations are present in all RIN stimuli, the subtraction will remove contributions from slow modulations. The response morphology for induced high gamma shows a burst of power from 70 to approximately 300 ms across subjects, and subsequently, in some electrode contacts, high gamma activity is found throughout the duration of the RIN. The induced high gamma response at 70 ms precedes the response that would be expected to spectro-temporal modulations at 4–8 Hz. In general, we found that subtraction of the 16-Hz response from higher-rate responses reduced the magnitude of high gamma responses more after 500 ms than it did between 70 and 300 ms (Supplementary Fig. 5), suggesting that the response prior to 300 ms is free of modulation contamination. Across all conditions, the difference in the proportion of significant electrode contact high gamma responses after subtraction of the 16 Hz RIN (compared to the noise period subtraction) was a decrease in 15.1% for the 0–500 ms analysis window, a decrease of 11.7% for 500–1000 ms, and a decrease of 30.9% for 1000–1500 ms (Supplementary Table 2).

3.4. Mapping of activity associated with RIN in the STP

To facilitate the comparison of HG induced responses across subjects, warped contacts coordinates were transferred and visualized on the CIT168 template brain (which is coregistered to the ICBM152 brain), separately for the left and right hemispheres. Fig. 4 shows an axial view with electrode contact locations on the STP. The response at a given electrode contact was scaled so that the permutation-test significance threshold corresponds to a value of 1, as observed for the 256 Hz RIN condition in the band from 70 to 120 Hz, when comparing the 500 ms preceding the transition from noise to RIN, to the 500 ms following the transition from noise to RIN. Electrode trajectories in the template brain were in good correspondence with the reconstructed locations in individual brains (see Supplementary Table 1 for MNI coordinates of each contact). For example, contacts that were anterior to the crest of HG on the subject's reconstructed image would be similarly located anterior to the crest of HG on the template brain. The combined responses in each hemisphere show a distribution of responses along the length of HG. Robust responses were also observed in PT from the one electrode sampling this region in the left hemisphere. Supplementary Fig. 6 shows the pattern of responses across all stimulus conditions using the same display as Fig. 4. The response distribution across the pitch-associated stimulus conditions are very similar and differ mostly in response magnitude as the strength of the pitch percept increases.

3.5. Mapping of activity associated with RIN to lateral temporal lobe

Responses were simultaneously obtained from electrodes covering portions of the lateral temporal lobe surface with recordings from 96-contact grids in each subject. Fig. 5 displays the lateral surface reconstructions for each subject along with lateral temporal electrode grid contact locations for each stimulus condition. Responses (70–120 Hz) at the contacts are thresholded for significance for each RIN condition from the comparison of the 500 ms preceding the transition from noise to RIN to the 500 ms immediately following the transition from noise to RIN. Significant responses ($p < 0.05$) are shown scaled to each subject's permutation threshold, as for Fig. 4. In each subject the induced response to the strongest pitch stimulus was a contiguous cluster of contacts (4–37) in the vicinity of transverse temporal sulcus.

4. Discussion

We recorded LFPs from human auditory cortex while participants listened to RIN stimuli with repetition rates below and above the LLP. The autocorrelation function of the LFP peaked at the RIN rate for rates above *and* below the LLP. Conversely, the induced high gamma response after the sound complex transition was effectively restricted to RIN conditions with rates above the LLP. The results indicate that the induced response is a better candidate for the neural correlate of pitch-related processing for repetition rate stimuli above the LLP than the general temporal regularity feature provided by autocorrelation. The evoked and

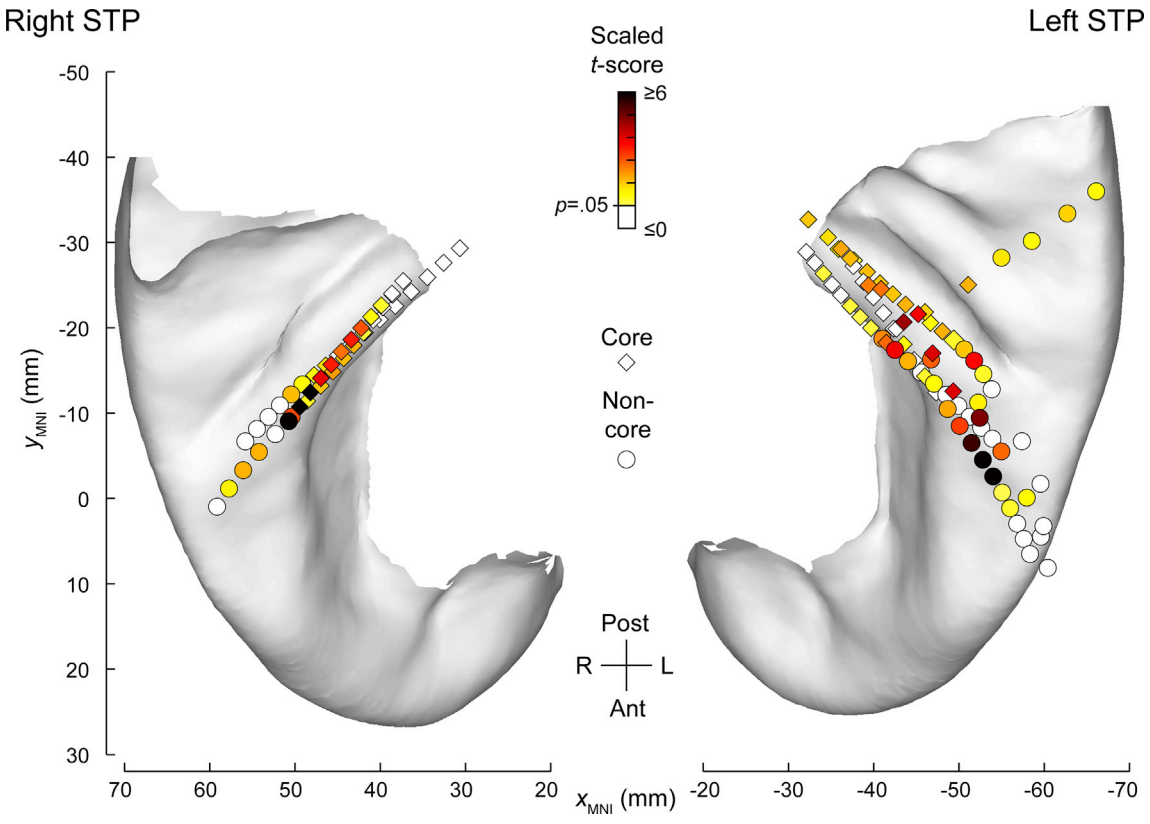

Fig. 4. STP high gamma response common mapping on template brain. Electrode locations are displayed on the STP surface of the ICBM152 template brain for the three right and six left STP depth electrodes (see Supplementary Table 1 for MNI coordinates). The view is top-down on the axial plane with all contacts shown on the surface for display purposes. The electrode symbols differentiate core and non-core auditory cortex contact locations determined for each subject's FFR to 100 Hz click train stimuli. The electrode contacts are not collinear because of co-registration of positions with preimplantation images, and reflecting distortion of anatomy due to electrode insertion. Responses are shown for the 256 Hz RIN in the band 70–120 Hz, from the comparison of the 500 ms preceding the transition from noise to RIN, to the 500 ms immediately following the transition from noise to RIN. Values are shown as scaled *t*-scores, with each individual subject's significance threshold being 1; responses greater than 6 are black and less than 1 are white, and negative values are indicated as zero. The black line in the scale bar marks the significance threshold determined from a permutation of *t*-scores scaled to the maximum for each subject. Responses are shown as the colored symbols at each contact location. Post = posterior, Ant = anterior, R = right, L = left.

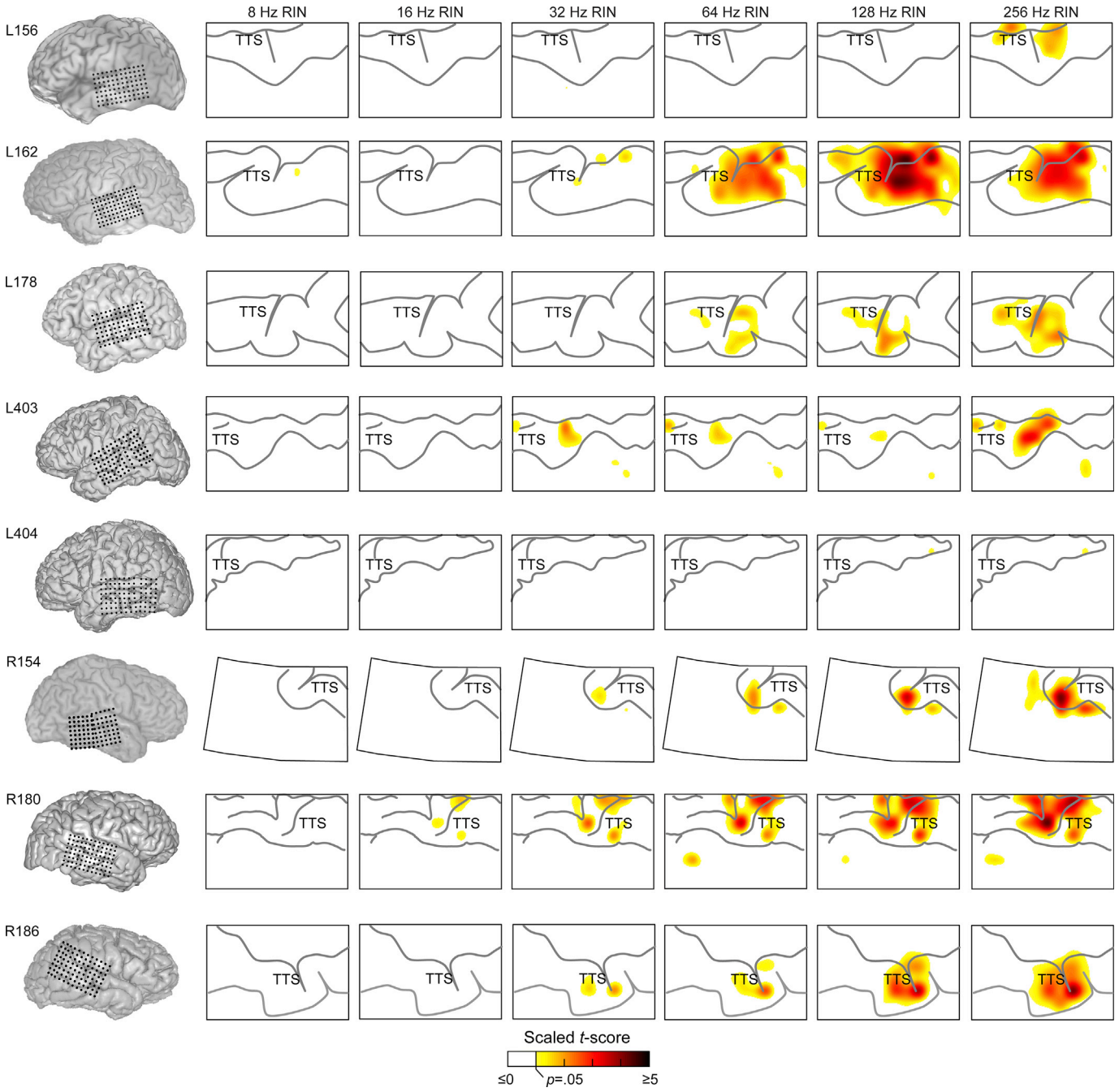

Fig. 5. High gamma ERSP on 96-contact lateral temporal grid electrode. Lateral surface reconstructions are shown for each subject with grids displayed. Responses (70–120 Hz) are shown for all stimulus conditions from the comparison of 500 ms following the transition from noise to RIN, to the 500 ms immediately preceding the transition from noise to RIN. Values are shown as scaled *t*-scores, with each individual subject's significance threshold being 1; responses greater than 5 are black and less than 1 are white, and negative values are indicated as zero. The smoothing applied to the data approximated 2 mm. TTS = transverse temporal sulcus.

induced responses have distinct distributions across the STP (see Fig. 2). High gamma induced responses to RIN are present in human core homologues of medial HG, in belt homologues of lateral HG, and belt/parabelt homologues in PT. In addition, high gamma responses to RIN were present in a region of the temporal convexity near the lateral extremity of HG. Importantly, high gamma responses are not being proposed as having the single property of representing pitch information, as high gamma responses are a reflection of local cortical processing (see further below). Instead, what is shown is the existence of high gamma activity, which is selectively expressed at repetition rates that correspond to the emergence of the perception of pitch, in particular auditory cortical regions.

The data reported here were recorded from patients with temporal lobe epilepsy, who have been found to show impairments in central auditory processing (Han et al., 2011), and may experience cortical reorganization due to the chronic nature of the disorder, which can influence normal brain function, for example language organization (Hamberger and Cole, 2011). Time limitations on recordings with patients allowed us to only assess a limited number of repetition rates around the LLP, despite the likelihood that phase locking to pure tones in humans likely occurs for frequencies around 1 kHz and possibly higher, and that perception of pitch for harmonic complexes may extend to around 8 kHz (Oxenham, 2018).

4.1. Evoked and induced activity due to RIN

Two previous studies have examined RIN evoked potential activity in the superior temporal plane and one study has examined induced activity. One study, using a noise-to-RIN transition in a single patient with an electrode implanted in PT parallel to HG, found a significant evoked response to pitch onset in the most lateral contact in PT (Schonweisner and Zatorre, 2008). Induced responses were not measured in that study. The low impedance clinical electrodes used in that study were estimated to be sensitive to LFP activity within a 1 cm radius providing a sampling of most of HG and much of PT. An earlier MEG study (Krumbholz et al., 2003) measured evoked responses to regular interval noise that were localized to medial HG. In the present study we observed evoked pitch onset responses measured by autocorrelation throughout HG with a general trend to responses being stronger in the posteromedial portion of HG. These results with eight subjects replicate and extend an earlier intracranial HG electrophysiology study in two subjects (Griffiths et al., 2010). That study showed that autocorrelation of the LFP increased with the regularity of the stimulus, but showed no clear relationship to the RIN repetition rate. The responses we demonstrate in the current report are similarly distributed to the evoked responses to click trains reported by Brugge et al. (2009), with a preference for responses in the medial half of HG and for lower repetition rates. Brugge et al. (2009) have proposed such phase-locking as a defining characteristic of human core homologues.

In our previous report (Griffiths et al., 2010), the high gamma induced responses occurred in more regular stimuli but were only present for repetition rates at or above the LLP. This study also found a direct relationship between the high gamma response and stimulus regularity (increase in iterations of RIN) for a fixed 128 Hz RIN. As the number of iterations of the delay-and-add cycle used to create a RIN stimulus are increased there is an increase in the perceptual salience of the pitch evoking sound (Yost, 1996). Both manipulations of RIN (repetition rate and iteration number) in the previous report changed the pitch salience. These high gamma responses to repetition rate and iteration number were observed in the same electrode contacts in both medial and lateral areas of HG. However, with recordings from only two subjects, no conclusions were drawn regarding the spatial distribution of pitch responses. The induced responses occurred throughout the entire gamma range in both subjects with a concentration of energy from 80 to 120 Hz. The present data provide further confirmation of the distribution of RIN-induced gamma responses across medial and lateral HG. Notably, the 64 and 128 Hz RIN conditions fully overlap the frequency range of the observed induced high gamma responses, and phase-locking to the repetition rate of harmonic complex tones has been observed up to F0s of 300 Hz in macaque primary auditory cortex (Fishman et al., 2013). However, the induced responses we observed are not just a reflection of the evoked responses phase locked to the RIN stimulus rates (or their sub-harmonics). This can be observed in power above and below the rates that overlap the high gamma range (compare contacts between Supplementary Fig. 1 and Supplementary Fig. 3).

The differing pattern of induced and evoked responses across the above studies has been suggested to be due to differences in neural generators for these response types (Plack et al., 2014). In comparison to the previous reports the present data in eight subjects demonstrate a broader distribution of responses sensitive to pitch from the evoked pitch-onset and induced high gamma measures that includes medial and lateral HG, PT, and lateral STG. Importantly however, the evoked and induced responses do show different distributional patterns, which is suggestive of different neural generators for these responses. Evoked responses have energy concentrated in a low frequency range (around 5–15 Hz) and are likely generated as the result of phase-locked activity of large groups of excitatory post-synaptic potentials on dendrites in cortical layer IV (Metherate and Cruikshank, 1999) but are also observable in other laminae (Steinschneider et al., 2008). Induced responses capture non-phase-locked activity, and in the high gamma frequency range likely represent increases in neuronal synchrony (Ray et al., 2008) and increases in multiunit spiking activity of large groups of pyramidal neurons in superficial layers (Ray and Maunsell, 2011; Bastos et al., 2012) that is the integration of activity and information from local cortical circuits (Traub et al., 2005; Crone et al., 2006; Tallon-Baudry, 2009). The different responses may represent separate aspects of a communication system across neuronal assemblies (Bastos et al., 2012) for prediction about pitch (Kumar et al., 2011).

4.2. Basis of the response to RIN

Recent fMRI studies have considered properties of the RIN stimulus other than pitch as an explanation for the associated brain activity. The delay-and-add procedure used to construct RIN also introduces slow spectral-temporal fluctuations over hundreds of milliseconds that are associated with changes in timbre rather than pitch, and are not present in a control noise (Barker et al., 2012; Hall and Plack, 2009; Plack et al., 2014). These fluctuations are present above and below the LLP. In the present work we carried out a control comparison (see Supplementary Fig. 5 and Supplementary Table 2) in which we subtract the response to 16 Hz RIN that is below the LLP: this stimulus contains the slow fluctuations described above. The induced responses to pitch are also present in the majority of contacts in this control condition that takes account of the timbral effect of the delay-and-add manipulation. Moreover, the gamma band responses latencies around 70 ms are too early to reflect the slow fluctuations that arise over hundreds of milliseconds. In general, subjects show an increasing function of induced gamma power as a function of stimulus rate across HG electrode contacts with a pronounced increase around the LLP (~30 Hz) (see Fig. 2). This pattern of results shows that the brain response tracks the perceptual properties of the pitch of the stimulus, while the slow spectro-temporal modulations are similar across all delay values.

An alternative explanation for the high gamma responses observed for RIN above the LLP in the present study may be based on the spectral resolvability of the stimuli. As the repetition rate of RIN stimuli increases, the spectral spacing and resolvability between the harmonics of the stimuli also increase. Because the stimuli were high pass filtered at 800 Hz, repetition rates at the higher frequencies have more resolved harmonics than the lower repetition rates (or are possibly only resolved for 128 and 256 Hz RIN). Recordings in macaque primary auditory cortex have demonstrated that neural population activity (multiunit and LFPs) increases as a function of the spacing and resolvability between harmonics of a harmonic complex tone (Fishman et al., 2000; Fishman and Steinschneider, 2006). Larger responses in anterolateral HG were found to resolved compared to unresolved harmonics of complex tones in a human fMRI study (Penagos et al., 2004).

4.3. Distribution of induced responses to RIN

The use of data from eight subjects allows us to investigate the distribution of induced responses to RIN that are within the cortex of the STP. Even with eight subjects, however, the data are far from a uniform sample and the interpretation of response distributions needs to be carried out with caution. The combined data for the group in Fig. 4 highlights the distribution of individual induced responses to RIN in both medial and lateral HG and PT (also shown in Figs. 2 and 3). There is no clearly discernible pattern with respect to hemisphere or the core/non-core boundary. Supplementary Fig. 6 shows the STP response distribution across the different stimuli, which shows a clear emergence of activity above the LLP that is similarly distributed across the pitch-evoking stimuli and whose response magnitude scales with pitch repetition rate.

The present study also demonstrates induced responses to RIN over the convexity of the superior temporal lobe where the use of grid arrays allows us to better claim a uniform (if sparse) sampling of the responses. The data in Fig. 5 show responses that are largely restricted to a limited region in the STG immediately adjacent to the transverse temporal

sulcus. In terms of homologues of primate brain, this area on the lateral convexity would be considered auditory parabelt in regions labelled RPB and/or CPB in monkey models. Notably however, these regions exhibit feedforward and feedback connections with auditory core (A1, R, RT) in tracer injection studies (Scott et al., 2017). Recent human invasive recording studies have identified this region to have response properties consistent with a relatively early processing stage within the auditory cortical hierarchy (i.e., likely at least in part represent the human homologue of lateral belt). These properties include: robust responses to pure tones during passive listening (Nourski et al., 2014a), reliable phase-locking to 100 Hz click trains (Nourski et al., 2013), short onset response latencies to acoustic stimuli (Nourski et al., 2014b), and very short latencies to electrical stimulation of A1, suggesting a monosynaptic connection (Brugge et al., 2003). The present data are supportive of this interpretation, showing a similarity of pitch responses with core cortex in posteromedial HG.

4.4. Comparison with human functional imaging studies of pitch

Gamma band responses above 30 Hz are known to correlate positively with fMRI BOLD (Mukamel et al., 2005) creating a potential link between studies using EEG, MEG, and ECoG with those of fMRI. However, the relationship between depth electrode recordings and fMRI BOLD is still not fully understood (Oya et al., 2018). A review of fMRI studies investigating human pitch processing by Plack et al. (2014) points to a possible area of specialized pitch processing in the region of lateral HG and adjacent anterolateral PT, which may differ across individuals. The present data show a response in lateral HG, however it is on the anterior side and no electrodes sampled posterolateral HG, or immediately adjacent anterolateral PT (see Fig. 4). Importantly, the present data show a response to pitch throughout the length of HG, that is not concentrated in one region, which, despite the lack of sampling in the present study in the region of anterolateral PT, precludes the possibility of an exclusive pitch processing region confined to that area. A distributed response along HG was also found in the RIN study of pitch by Patterson et al. (2002), with the strongest response found in lateral HG. A recent fMRI study investigating pitch responses and their relationship to harmonic resolvability, behavioral thresholds, and tonotopy found pitch sensitive responses that spanned anterolateral PT, lateral HG, and lateral planum polare (Norman-Haignere et al., 2013). This distribution of responses includes the location of the anterolateral HG clusters in the present data – the region with the largest number of subjects with pitch sensitive voxels in lateral HG in their report. While not included as an ROI in the pitch responsive regions of Norman-Haignere et al. (2013), the location of the response on the lateral STG convexity surrounding transverse temporal sulcus in the current results is also in this same approximate location. However, unlike in their data, no clear pattern along HG of increased responsiveness in a gradient from posterior to anterior was observed in the present data set. A recent meta-analysis of auditory fMRI studies of pitch processing found the median activity locus to be at or near PT in studies with infrequent pitch changes in tone streams (Alho et al., 2014), a different stimulus than employed in the present report. Activity for passive and active listening to tones with fixed or varying pitch was found on STG lateral to HG, with a tendency for the median locus to be near transverse temporal sulcus, similar to the stimulus and results in the current report.

4.5. Comparison with animal models for pitch perception

The human imaging studies based on measures of neuronal ensemble activity above have suggested a number of different loci for pitch analysis that are largely in non-core areas. Similarly, neural population responses representing the pitch of missing fundamental harmonic complexes were found to not likely exist in auditory core in macaque monkeys (Fishman et al., 1998). Neuronal recordings in the ferret have demonstrated responses to pitch that are distributed across core and belt areas (Walker et al., 2011), while neuronal recordings in the marmoset using missing-fundamental stimuli associated with pitch demonstrated only a small region of pitch-selective neurons in an area overlapping auditory core and anterolateral to auditory core (Bendor and Wang, 2005, 2010). More recent findings in the marmoset indicate there is a unique population of neurons distributed across core auditory cortex that selectively responds to harmonically structured sounds (Feng and Wang, 2017). Whether this population of neurons accounts for the present findings is unclear. Recent work on the macaque based on both single-neuron and LFP recordings to both harmonic pitch and RIN demonstrated distributed responses to RIN pitch in both core and non-core auditory cortex (Kikuchi et al., in press).

Comparison of pitch mechanisms across animal species is problematic, even amongst primates [numbering over 500 species; Estrada et al. (2017)], which show different auditory cortical organization in New World species (like marmoset) with auditory cortex on the exterior temporal lobe, and Old World species (like macaque) with auditory cortex within a superior temporal plane. The closest widely available model for human auditory cortex is the macaque, but even in that species there is controversy about the basis for the homology (Baumann et al., 2013; Norman-Haignere et al., 2019). In the present study we demonstrate responses to differing repetition rates that are parsimoniously explained as pitch-related responses because they respect a perceptual boundary, that are distributed across both core and non-core homologues. Future work on the computational code for pitch to determine necessary and sufficient conditions of pitch processing will need to demonstrate behavioral relevance of the neural activity as it relates to perception (Wang and Walker, 2012). The data are not consistent with a single pitch center, but rather demonstrate the architecture of a system distributed across auditory cortex for human pitch processing. The similarity of pitch-related responses in the high gamma range across multiple auditory cortical regions may indicate that a similar code is shared across these regions for a single pitch-related feature (e.g., stimulus repetition rate). Alternatively, a distributed processing system for pitch may exist such that different levels of the auditory pathway are responsible for the representation of particular features used to create pitch and that communication among these regions occurs to construct a percept based on feedback to areas earlier in the computational hierarchy (Kumar et al., 2011).

Acknowledgements

This work was supported by the NIH (DC004290, 1S10OD025025-01), the Wellcome Trust (WT091681MA), and the Hoover Fund. This work was conducted on an MRI instrument funded by 1S10OD025025-01. We thank Fangxiang Chen, Haiming Chen, Richard Reale, and Chandan Reddy for their help with conducting the experiments. We would like to thank the patients who contributed their time for the experiments. The authors declare no competing financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.116076.

References

Alho, K., Rinne, T., Herron, T.J., Woods, D.L., 2014. Stimulus-dependent activations and attention-related modulations in the auditory cortex: a meta-analysis of fMRI studies. Hear. Res. 307, 29–41.

Avants, B.B., Duda, J.T., Zhang, H., Gee, J.C., 2007. Multivariate normalization with symmetric diffeomorphisms for multivariate studies. Med Image Comput Comput Assist Interv 10, 359–366.

Barker, D., Plack, C.J., Hall, D.A., 2012. Reexamining the evidence for a pitch-sensitive region: a human fMRI study using iterated ripple noise. Cerebr. Cortex 22, 745–753.

Barrett, D.J., Hall, D.A., 2006. Response preferences for "what" and "where" in human non-primary auditory cortex. Neuroimage 32, 968–977.

Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., Friston, K.J., 2012. Canonical microcircuits for predictive coding. Neuron 76, 695–711.

Baumann, S., Petkov, C.I., Griffiths, T.D., 2013. A unified framework for the organization of the primate auditory cortex. Front. Syst. Neurosci. 7, 11.
Bendor, D., Wang, X., 2005. The neuronal representation of pitch in primate auditory cortex. Nature 436, 1161–1165.
Bendor, D., Wang, X., 2010. Neural coding of periodicity in marmoset auditory cortex. J. Neurophysiol. 103, 1809–1822.
Bizley, J.K., Walker, K.M., King, A.J., Schnupp, J.W., 2010. Neural ensemble codes for stimulus periodicity in auditory cortex. J. Neurosci. 30, 5078–5091.
Bizley, J.K., Walker, K.M., Nodal, F.R., King, A.J., Schnupp, J.W., 2013. Auditory cortex represents both pitch judgments and the corresponding acoustic cues. Curr. Biol. 23, 620–625.
Brugge, J.F., Volkov, I.O., Garell, P.C., Reale, R.A., Howard 3rd, M.A., 2003. Functional connections between auditory cortex on Heschl's gyrus and on the lateral superior temporal gyrus in humans. J. Neurophysiol. 90, 3750–3763.
Brugge, J.F., Nourski, K.V., Oya, H., Reale, R.A., Kawasaki, H., Steinschneider, M., Howard 3rd, M.A., 2009. Coding of repetitive transients by auditory cortex on Heschl's gyrus. J. Neurophysiol. 102, 2358–2374.
Cedolin, L., Delgutte, B., 2010. Spatiotemporal representation of the pitch of harmonic complex tones in the auditory nerve. J. Neurosci. 30, 12712–12724.
Conway, B.R., 2009. Color vision, cones, and color-coding in the cortex. Neuroscientist 15, 274–290.
Crone, N.E., Sinai, A., Korzeniewska, A., 2006. High-frequency gamma oscillations and human brain mapping with electrocorticography. Prog. Brain Res. 159, 275–295.
Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.
Estrada, A., Garber, P.A., Rylands, A.B., Roos, C., Fernandez-Duque, E., Di Fiore, A., Nekaris, K.A., Nijman, V., Heymann, E.W., Lambert, J.E., Rovero, F., Barelli, C., Setchell, J.M., Gillespie, T.R., Mittermeier, R.A., Arregoitia, L.V., de Guinea, M., Gouveia, S., Dobrovolski, R., Shanee, S., Shanee, N., Boyle, S.A., Fuentes, A., MacKinnon, K.C., Amato, K.R., Meyer, A.L., Wich, S., Sussman, R.W., Pan, R., Kone, I., Li, B., 2017. Impending extinction crisis of the world's primates: why primates matter. Sci Adv 3, e1600946.
Feng, L., Wang, X., 2017. Harmonic template neurons in primate auditory cortex underlying complex sound processing. Proc. Natl. Acad. Sci. U. S. A. 114, E840–E848.
Fishman, Y.I., Steinschneider, M.J., 2006. Spectral resolution of monkey primary auditory cortex (A1) revealed with two-noise masking. Neurophysiology 96, 1105–1115.
Fishman, Y.I., Reser, D.H., Arezzo, J.C., Steinschneider, M., 2000. Complex tone processing in primary auditory cortex of the awake monkey. II. Pitch versus critical band representation. J. Acoust. Soc. Am. 108, 247–262.
Fishman, Y.I., Micheyl, C., Steinschneider, M., 2013. Neural representation of harmonic complex tones in primary auditory cortex of the awake monkey. J. Neurosci. 33, 10312–10323.
Fishman, Y.I., Reser, D.H., Arezzo, J.C., Steinschneider, M., 1998. Pitch vs. spectral encoding of harmonic complex tones in primary auditory cortex of the awake monkey. Brain Res. 786, 18–30.
Fonov, V., Evans, A.C., Botteron, K., Almli, C.R., McKinstry, R.C., Collins, D.L., Brain Development Cooperative, G., 2011. Unbiased average age-appropriate atlases for pediatric studies. Neuroimage 54, 313–327.
Galaburda, A., Sanides, F., 1980. Cytoarchitectonic organization of the human auditory cortex. J. Comp. Neurol. 190, 597–610.
Galambos, R., Makeig, S., Talmachoff, P.J., 1981. A 40-Hz auditory potential recorded from the human scalp. Proc. Natl. Acad. Sci. U. S. A. 78, 2643–2647.
Griffiths, T.D., Hall, D.A., 2012. Mapping pitch representation in neural ensembles with fMRI. J. Neurosci. 32, 13343–13347.
Griffiths, T.D., Kumar, S., Sedley, W., Nourski, K.V., Kawasaki, H., Oya, H., Patterson, R.D., Brugge, J.F., Howard, M.A., 2010. Direct recordings of pitch responses from human auditory cortex. Curr. Biol. 20, 1128–1132.
Hackett, T.A., Preuss, T.M., Kaas, J.H., 2001. Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. J. Comp. Neurol. 441, 197–222.
Hall, D.A., Plack, C.J., 2009. Pitch processing sites in the human auditory brain. Cerebr. Cortex 19, 576–585.
Hamberger, M.J., Cole, J., 2011. Language organization and reorganization in epilepsy. Neuropsychol. Rev. 21, 240–251.
Han, M.W., Ahn, J.H., Kang, J.K., Lee, E.M., Lee, J.H., Bae, J.H., Chung, J.W., 2011. Central auditory processing impairment in patients with temporal lobe epilepsy. Epilepsy Behav. 20, 370–374.
Howard 3rd, M.A., Volkov, I.O., Granner, M.A., Damasio, H.M., Ollendieck, M.C., Bakken, H.E., 1996. A hybrid clinical-research depth electrode for acute and chronic in vivo microelectrode recording of human brain neurons. Technical note. J. Neurosurg. 84, 129–132.
Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. Fsl. Neuroimage 62, 782–790.
Joly, O., Baumann, S., Poirier, C., Patterson, R.D., Thiele, A., Griffiths, T.D., 2014. A perceptual pitch boundary in a non-human primate. Front. Psychol. 5, 998.
Kikuchi, Y., Kumar, S., Baumann, S., Overath, T., Gander, P.E., Sedley, W., Patterson, R., Petkov, C.I., Griffiths, T.D., (in press). The distribution and nature of responses to broadband sounds associated with pitch in the macaque auditory cortex. Cortex.
Kovach, C.K., Gander, P.E., 2016. The demodulated band transform. J. Neurosci. Methods 261, 135–154.
Krumbholz, K., Patterson, R.D., Pressnitzer, D., 2000. The lower limit of pitch as determined by rate discrimination. J. Acoust. Soc. Am. 108, 1170–1180.
Krumbholz, K., Patterson, R.D., Seither-Preisler, A., Lammertmann, C., Lutkenhoner, B., 2003. Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. Cerebr. Cortex 13, 765–772.
Kumar, S., Sedley, W., Nourski, K.V., Kawasaki, H., Oya, H., Patterson, R.D., Howard 3rd, M.A., Friston, K.J., Griffiths, T.D., 2011. Predictive coding and pitch processing in the auditory cortex. J. Cogn. Neurosci. 23, 3084–3094.
Liegeois-Chauvel, C., Musolino, A., Chauvel, P., 1991. Localization of the primary auditory area in man. Brain 114 (Pt 1A), 139–151.
Metherate, R., Cruikshank, S.J., 1999. Thalamocortical inputs trigger a propagating envelope of gamma-band activity in auditory cortex in vitro. Exp. Brain Res. 126, 160–174.
Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., Zilles, K., 2001. Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage 13, 684–701.
Mukamel, R., Gelbard, H., Arieli, A., Hasson, U., Fried, I., Malach, R., 2005. Coupling between neuronal firing, field potentials, and FMRI in human auditory cortex. Science 309, 951–954.
Nagahama, Y., Kovach, C.K., Ciliberto, M., Joshi, C., Rhone, A.E., Vesole, A., Gander, P.E., Nourski, K.V., Oya, H., Howard 3rd, M.A., Kawasaki, H., Dlouhy, B.J., 2018. Localization of musicogenic epilepsy to Heschl's gyrus and superior temporal plane: case report. J. Neurosurg. 129, 157–164.
Norman-Haignere, S., Kanwisher, N., McDermott, J.H., 2013. Cortical pitch regions in humans respond primarily to resolved harmonics and are located in specific tonotopic regions of anterior auditory cortex. J. Neurosci. 33, 19451–19469.
Nourski, K.V., Brugge, J.F., Reale, R.A., Kovach, C.K., Oya, H., Kawasaki, H., Jenison, R.L., Howard 3rd, M.A., 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, K.V., Steinschneider, M., Oya, H., Kawasaki, H., Jones, R.D., Howard, M.A., 2014a. Spectral organization of the human lateral superior temporal gyrus revealed by intracranial recordings. Cerebr. Cortex 24, 340–352.
Nourski, K.V., Steinschneider, M., McMurray, B., Kovach, C.K., Oya, H., Kawasaki, H., Howard 3rd, M.A., 2014b. Functional organization of human auditory cortex: investigation of response latencies through direct recordings. Neuroimage 101, 598–609.
Nourski, K.V., Steinschneider, M., Rhone, A.E., 2016. Electrocorticographic activation within human auditory cortex during dialog-based language and cognitive testing. Front. Hum. Neurosci. 10, 202.
Oxenham, A.J., 2018. How we hear: the perception and neural coding of sound. Annu. Rev. Psychol. 69, 27–50.
Oya, H., Kawasaki, H., Nourski, K.V., Greenlee, J.D., Griffiths, T.D., Howard III., M.A., 2014. Mapping intracranial ECoG responses onto the surface of the human superior temporal plane. In: Society for Neuroscience 44th Annual Meeting. November 15-19, 2014, Washington, DC. Program No. 235.05. 2014 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience, 2014. Online.
Oya, H., Gander, P.E., Petkov, C.I., Adolphs, R., Nourski, K.V., Kawasaki, H., Howard, M.A., Griffiths, T.D., 2018. Neural phase locking predicts BOLD response in human auditory cortex. Neuroimage 169, 286–301.
Patterson, R.D., Uppenkamp, S., Johnsrude, I.S., Griffiths, T.D., 2002. The processing of temporal pitch and melody information in auditory cortex. Neuron 36, 767–776.
Penagos, H., Melcher, J.R., Oxenham, A.J., 2004. A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. J. Neurosci. 24, 6810–6815.
Plack, C.J., Barker, D., Hall, D.A., 2014. Pitch coding and pitch processing in the human brain. Hear. Res. 307, 53–64.
Pressnitzer, D., Patterson, R.D., Krumbholz, K., 2001. The lower limit of melodic pitch. J. Acoust. Soc. Am. 109, 2074–2084.
Puschmann, S., Uppenkamp, S., Kollmeier, B., Thiel, C.M., 2010. Dichotic pitch activates pitch processing centre in Heschl's gyrus. Neuroimage 49, 1641–1649.
Ray, S., Crone, N.E., Niebur, E., Franaszczuk, P.J., Hsiao, S.S., 2008. Neural correlates of high-gamma oscillations (60-200 Hz) in macaque local field potentials and their potential implications in electrocorticography. J. Neurosci. 28, 11526–11536.
Ray, S., Maunsell, J.H., 2011. Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. PLoS Biol. 9, e1000610.
Reddy, C.G., Dahdaleh, N.S., Albert, G., Chen, F., Hansen, D., Nourski, K., Kawasaki, H., Oya, H., Howard 3rd, M.A., 2010. A method for placing Heschl gyrus depth electrodes. J. Neurosurg. 112, 1301–1307.
Reuter, M., Rosas, H.D., Fischl, B., 2010. Highly accurate inverse consistent registration: a robust approach. Neuroimage 53, 1181–1196.
Ross, B., Borgmann, C., Draganova, R., Roberts, L.E., Pantev, C., 2000. A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. J. Acoust. Soc. Am. 108, 679–691.
Schonwiesner, M., Zatorre, R.J., 2008. Depth electrode recordings show double dissociation between pitch processing in lateral Heschl's gyrus and sound onset processing in medial Heschl's gyrus. Exp. Brain Res. 187, 97–105.
Scott, B.H., Leccese, P.A., Saleem, K.S., Kikuchi, Y., Mullarkey, M.P., Fukushima, M., Mishkin, M., Saunders, R.C., 2017. Intrinsic connections of the core auditory cortical regions and rostral supratemporal plane in the macaque monkey. Cerebr. Cortex 27, 809–840.
Steinschneider, M., Fishman, Y.I., Arezzo, J.C., 2008. Spectrotemporal analysis of evoked and induced electroencephalographic responses in primary auditory cortex (A1) of the awake monkey. Cerebr. Cortex 18, 610–625.
Talavage, T.M., Ledden, P.J., Benson, R.R., Rosen, B.R., Melcher, J.R., 2000. Frequency-dependent responses exhibited by multiple regions in human auditory cortex. Hear. Res. 150, 225–244.
Tallon-Baudry, C., 2009. The roles of gamma-band oscillatory synchrony in human visual cognition. Front. Biosci. 14, 321–332.

Traub, R.D., Bibbig, A., LeBeau, F.E., Cunningham, M.O., Whittington, M.A., 2005. Persistent gamma oscillations in superficial layers of rat auditory neocortex: experiment and model. J. Physiol. 562, 3–8.
Tyszka, J.M., Pauli, W.M., 2016. In vivo delineation of subdivisions of the human amygdaloid complex in a high-resolution group template. Hum. Brain Mapp. 37, 3979–3998.
Walker, K.M., Bizley, J.K., King, A.J., Schnupp, J.W., 2011. Multiplexed and robust representations of sound features in auditory cortex. J. Neurosci. 31, 14565–14576.
Wang, X., Walker, K.M., 2012. Neural mechanisms for the abstraction and use of pitch information in auditory cortex. J. Neurosci. 32, 13339–13342.
Yeo, B.T., Sabuncu, M.R., Vercauteren, T., Ayache, N., Fischl, B., Golland, P., 2010. Spherical demons: fast diffeomorphic landmark-free surface registration. IEEE Trans. Med. Imaging 29, 650–668.
Yost, W., 1996. Pitch of iterated rippled noise. J. Acoust. Soc. Am. 100, 511–518.