NeuroImage 238 (2021) 118222

Contents lists available at ScienceDirect

NeuroImage

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

Mapping the human auditory cortex using spectrotemporal receptive fields generated with magnetoencephalography

Jean-Pierre R. Falet[1], Jonathan Côté[1,*], Veronica Tarka, Zaida Escila Martínez-Moreno, Patrice Voss, Etienne de Villers-Sidani

Department of Neurology and Neurosurgery, McGill University, Canada

ARTICLE INFO

Keywords:
Magnetoencephalography
MEG
Tonotopy
Auditory cortex
Spectrotemporal receptive field
Abbreviations
IIPT
Isointensity Pure Tone

ABSTRACT

We present a novel method to map the functional organization of the human auditory cortex noninvasively using magnetoencephalography (MEG). More specifically, this method estimates via reverse correlation the spectrotemporal receptive fields (STRF) in response to a temporally dense pure tone stimulus, from which important spectrotemporal characteristics of neuronal processing can be extracted and mapped back onto the cortex surface. We show that several neuronal populations can be found examining the spectrotemporal characteristics of their STRFs, and demonstrate how these can be used to generate tonotopic gradient maps. In doing so, we show that the spatial resolution of MEG is sufficient to reliably extract important information about the spatial organization of the auditory cortex, while enabling the analysis of complex temporal dynamics of auditory processing such as best temporal modulation rate and response latency given its excellent temporal resolution. Furthermore, because spectrotemporally dense auditory stimuli can be used with MEG, the time required to acquire the necessary data to generate tonotopic maps is significantly less for MEG than for other neuroimaging tools that acquire BOLD-like signals.

1. Introduction

An important goal of auditory neurophysiology is to model the functional organization of the human auditory cortex (AC). This involves developing an understanding of auditory processing along both spectral and temporal dimensions, and relating these features to the spatial topographical organization of the AC.

Frequently, the topographical organization of the human AC has been studied noninvasively using functional magnetic resonance imaging (fMRI) in terms of tonotopy, or best frequency maps, which has been found to be a key organizational feature (Da Costa et al., 2011; Formisano et al., 2003; Humphries et al., 2010; Langers and van Dijk, 2012; Talavage and Edmister, 2004; Woods et al., 2010). Although details such as the orientation of the tonotopic gradient are still debated, an anterior to posterior high-low-high best frequency organization centered on Heschl's gyrus (HG) is found and agreed upon in most human fMRI studies (Gardumi et al., 2017), and is consistent whether pure tones or natural sounds are used (Moerel et al., 2012). Coupled with the spatial organization of other neuronal response characteristics such as the broadness of frequency tuning, and paired with findings from cyto- and myeloarchitectural studies, the AC has been further divided by fMRI into subfields with unique processing properties (Moerel et al., 2014).

However, the role of temporal processing within the microorganization of the human AC remains unclear from the available fMRI literature alone (Leaver and Rauschecker, 2016). Crucial aspects of our sensory experience, such as speech perception and music enjoyment, clearly rely heavily on precise temporal encoding of auditory information (Abrams et al., 2011). Invasive electrophysiological recordings in several animal species have shown the importance of temporal features in understanding the functionality of AC subfields (Linden et al., 2003; Nagel and Doupe, 2008). Specific areas have been reported to produce phasic or a tonic response to tones (Joachimsthaler et al., 2014). Additionally, auditory ERPs and single-unit recordings have shown different time courses as a function of stimuli durations (Beukes et al., 2009; Alain et al., 1997). Moreover, studying the temporal domain of auditory processing is necessary to gain a complete understanding of auditory plasticity (Schreiner and Polley, 2014; Carlin and Elhilali, 2015). For example, auditory training using temporal discrimination tasks can lead to improvements in the processing of temporal features that do not result in improvements in spectral processing (van Wassenhove and Nagarajan, 2007), reinforcing the importance of studying both dimen-

* Corresponding author.
E-mail addresses: jean-pierre.falet@mail.mcgill.ca (J.R. Falet), jonathan.cote@mail.mcgill.ca (J. Côté), veronica.tarka@mail.mcgill.ca (V. Tarka), zaida.martinezmoreno@mail.mcgill.ca (Z.E. Martínez-Moreno), patrice.voss@mcgill.ca (P. Voss), etienne.de-villers-sidani@mcgill.ca (E. de Villers-Sidani).
[1] Both the authors contributed equally to this work.

https://doi.org/10.1016/j.neuroimage.2021.118222.
Received 12 April 2021; Received in revised form 25 May 2021; Accepted 28 May 2021
Available online 29 May 2021.
1053-8119/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

sions. Similarly, studying temporal dynamics can yield insights into age-related changes in auditory processing (de Villers-Sidani et al., 2010; Ross et al., 2020; Dobri and Ross, 2021).

Unfortunately, while fMRI boasts an excellent spatial resolution to answer questions pertaining to the spatial organization of the AC, it cannot provide sufficient temporal resolution to adequately study temporal dynamics and short-latency events. The hemodynamic response to neuronal activity measured by fMRI occurs on the order of seconds (Aguirre et al., 1998), which precludes precise characterization of neuronal activity occurring on the order of milliseconds. Furthermore, because of the relatively long acquisition time, stimuli sets are typically small and offer less flexibility than one would ideally want to study the response to complex sounds. Studying auditory processing in fMRI has also been limited by loud operating noise, even though workarounds have been developed (Cha et al., 2016).

MEG is an attractive alternative modality for *in vivo* electrophysiological recording of neuronal activity in the AC. It not only provides superior temporal resolution on the order of milliseconds (Regan, 1989), but also provides a completely silent acquisition environment. An important barrier preventing its widespread use has been related to concerns regarding its ability to spatially resolve the millimetric spatial organization of the AC (Langers and van Dijk, 2012; Moerel et al., 2014), in particular its tonotopic organization. This concern is offset by recent successes in capturing the retinotopic organization of the visual cortex using MEG at a spatial resolution of 7 mm in smooth cortical regions and of less than 1 mm near curved gyri (Nasiotis et al., 2017). Furthermore, the MEG spatial resolution of the auditory cortex, given its dipole orientation, has been known to be under 1 cm (Pantev et al., 1990). Additionally, similarly fine cortical resolutions were obtained in other ROIs (Huang et al., 2016; Huang et al., 2006) and a human skull phantom (Leahy et al., 1998). A sufficiently precise resolution was demonstrated for preoperative source localization (Niranjan et al., 2013). Moreover, early efforts at identifying a basic tonotopic gradient using MEG have been successful in some respects: dipole depth beneath the scalp has consistently been found to correlate with stimulus frequency, and orientation of the gradient has been shown to vary with gyral morphology (Romani et al., 1982; Pantev et al., 1988; Kuriki and Murase, 1989; Huotilainen et al., 1995; Verkindt et al., 1995). Other studies have also identified a posterior to anterior gradient, lower frequencies being represented more posteriorly, with the possibility of there being multiple tonotopic gradients (Pantev et al., 1995; Weisz et al., 2004). Finally, a recent MEG study using speech sounds was able to identify a tonotopic gradient similar to that obtained in fMRI (Su et al., 2014).

Encouragingly, relatively simple study design tweaks could potentially yield improvements in the spatial resolution of MEG, notably through the use of higher stimulus density (ie. a higher stimulus presentation rate). There is evidence from research with owl monkeys pointing to an inverse relationship between stimulus density and the tuning width of neurons in the AC, as shown by the smaller size of their receptive fields with such stimuli (Blake and Merzenich, 2002). This could be due to increased peri-neuronal inhibition when stimuli are presented at a faster rate, increasing the spectrotemporal specificity of each neuron, and therefore improving the spatial resolvability of neuronal subpopulations. Using a temporally dense stimulus could therefore improve the spatial resolution of MEG with respect to tonotopic organization.

Here, we describe a novel method to map the functional organization of the AC using MEG. Specifically, we take advantage of the MEG's high temporal resolution to extract the spectral and temporal characteristics of sound processing for each neuronal source by computing their spectrotemporal receptive field (STRF), and demonstrate how the characteristics of STRFs can then be extracted and mapped onto the cortical surface to study organizational features such as tonotopy. STRFs have indeed been commonly used to describe the dynamics of neuronal activity in response to auditory stimuli (see for e.g.: Calabrese et al., 2011; Kowalski et al., 1996; Linden et al., 2003; Sen et al., 2001; Woolley et al., 2006). They represent the spectral and temporal patterns of auditory stimuli that elicit the maximal response from a neuron. To estimate STRFs, several methods have been used for varying stimulus types (Theunissen et al., 2000), but the foundational technique revolves around reverse correlation and involves averaging the stimulus content preceding neuronal spikes (de Boer and Kuyper, 1968). Doing so results in a spike-triggered average that can reliably estimate the STRF when using a stimulus that is uncorrelated in the spectral and temporal dimensions, as is typical for stimuli used for mapping tonotopy.

We show here that spectrotemporally dense auditory stimuli composed of isointensity pure tones (IIPTs) can yield sufficient spatial resolution to allow for mapping the tonotopic organization of the AC using reverse correlation-based STRFs generated from MEG. This method can therefore be reliably used to investigate the spatial organization of the AC, with the added benefit of MEG's excellent temporal resolution to study short-latency-dependent events and complex spectrotemporal characteristics, permitting an in-depth non-invasive functional study of auditory processing in humans.

2. Materials and methods

2.1. Participants

Ten right-handed participants were recruited into the study (henceforth labeled S1 to S10). Three were female and the average age was 23 (range 19–27). All participants reported being free of hearing impairment or neurological conditions that could affect brain function, including mild cognitive impairment, dementia and previous history of stroke. All participants provided written informed consent. This study was approved by the research ethics board of the Montreal Neurological Institute.

The MEG and anatomical MRI recordings of S3 are freely available for download from the OpenNeuro platform at the following link: https://openneuro.org/datasets/ds003082/versions/1.0.0 (Cote and de Villers-Sidani, 2020)

2.2. MEG analysis

2.2.1. Stimuli presentation

Stimuli were generated by a Sound Blaster X-Fi Titanium HD audio card (Creative, Jurong East, Singapore) connected to a pair of E-A-RTONE 3A insert earphones (3 M company, Indianapolis, Indiana). The earphones were connected to the inducers by a plastic tube of approximately 1 m in length. The inducers were tucked under a shielded sheathe on the floor. The audio output was compared to the generated stimulus with a sound level meter and an ear-canal adapter. Significant attenuation was measured at frequencies higher than 7 kHz, therefore all presentations above this value were not included in data analyses.

The stimulus train was a 10-minute train of 50-ms long gated IIPTs (Fig. 1). Each pure tone had an up and down-ramp of 5 ms. Thirty two different frequencies were presented with A-weighted intensities for the resulting stimulus train to be perceived at a similar intensity (Fletcher and Munson, 1933). More specifically, A-weighting describes the decibel attenuation necessary for each frequency to be perceived at the same intensity, because the perceived intensity varies depending on the frequency of the stimulus. There was an average of 55 pure tones per frequency, per recording, totalling 1795 pure tones per recording. This presentation rate and tone duration make the IIPT stimulus a non-verbal, pure tone-based close analog of the previously used Rapid Serial Auditory Presentation (Franco et al., 2015). Frequencies ranged from 0.1 kHz to 25.6 kHz, each separated by a quarter of an octave. The inter-stimulus interval was randomly generated from a gamma distribution with shape parameter 6 to achieve an average presentation rate of 3 Hz. Tones could overlap but less than 1% of tones did, and only 2 tones for every 64 were adjacent.

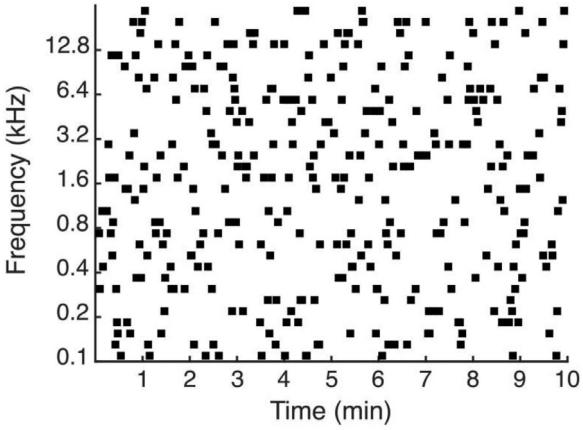

Fig. 1. Iso-intensity pure tones stimulus. Sample stimulus spectrogram used to obtain STRFs. Frequencies range from 0.1 kHz to 25.6 kHz and each is separated by a quarter of an octave. Tones are presented binaurally at an average rate of 3 Hz for a total of 10 min. Note that the width of the squares in this figure is larger than their true duration (50 ms) for the sake of presentation, and more tones appear to be overlapping than is truly the case. Refer to Fig. 2. A for a true representation of the duration of each stimulus over a smaller time-window.

Participants were instructed to fixate on a visual fixation cross throughout the stimulus presentation to reduce eye movement artifacts. The volume intensity was set to a comfortable hearing level.

2.2.2. MEG acquisition

Using a six-degrees-of-freedom digitizer (Patriot - Polhemus; Matlab interface RRID: SCR_006752) each participant's head was digitized. The head shapes contained about 100 to 200 points distributed across the scalp, eyebrows and nose to precisely coregister the activity to the structural MRI. Three coils were attached to fiducial anatomical locations on the head (nasion, and left and right pre-auricular points) to capture head movement inside the MEG. Localization was compared between the first and last second of recording to ensure no total movement larger than 1 cm. To record blinks and eye movements, we placed bipolar electro-oculographic (EOG) leads about 1 cm above and below one eye, and about 1 cm lateral of the outer canthi. Electrocardiographic (ECG) activity was recorded with one channel. The electrical reference was placed at the opposite clavicle. Both EOGs, ECG and the electrical reference were used for subsequent MEG artifact detection and removal. MEG was recorded using a 275-channel (axial gradiometers) whole-head MEG system (CTF MEG International Services Ltd.). All data were downsampled to 2400 Hz.

2.2.3. Structural MRI

Three-dimensional T1-weighted anatomical MR image volumes covering the entire brain were acquired on either a 1.5T Siemens Sonata or 3T Siemens Magnetom Prisma scanner with an 8-channel head coil (repetition time = 27 ms; echo time = 9.20 ms; between 176 and 192 sagittally oriented slices with slice thickness of 1 mm; acquisition matrix = 240 × 256; field of view = 256 mm).

2.2.4. MEG data pre-processing and spatial modeling

MEG data analysis was performed in Matlab (RRID: SCR_001622; MATLAB and Statistics Toolbox Release 2015b), coupled with the Brainstorm extension (Tadel et al., 2011), which is documented and freely available for download online under the GNU general public license (RRID: SCR_001761; Tadel, 2019). MRI-based cortical reconstruction and volumetric segmentation were performed with the FreeSurfer image analysis suite (RRID: SCR_001847; Fischl, 2013; Dale and Sereno, 1993; Fischl et al., 1999, 2001).

Raw MEG data were pre-processed to remove signal contamination due to ocular, cardiac, and muscular artifacts using signal-space projections (Tesche et al., 1995; Uusitalo and Ilmoniemi, 1997). The SSP projection was done in accordance with the Brainstorm tutorial's recommended procedure (Tadel, 2019), namely starting with the event detection of both eye blinks and heartbeats, followed by the removal of heartbeat temporally adjacent to eye blinks, calculation of SSP and the removal of heartbeat, and finally, calculation of the SSP and the removal of eye blinks. Each recording was then manually reviewed to discard any segment still experiencing significant contamination from artifacts. All data were downsampled from the recorded 12 kHz to 2400 Hz.

The forward problem was solved using the overlapping-sphere approach (Huang et al., 1999), which fits a sphere to the scalp surface. This simplified modeling method can be used given that the magnetic fields recorded from the brain are not distorted by the shape of the skull (Barth et al., 1986; Okada et al., 1999). wMNE (Lin et al., 2006) was used to solve the reverse problem, with sources being constrained to a one-dimensional perpendicular orientation with respect to the cortex surface. The MRI-based cortex surface was generated with FreeSurfer and contained 330,000 sources (Dale and Sereno, 1993). Otherwise, default Brainstorm parameters were used in the wMNE modeling (SNR: 3 / Whitening: PCA; Regularize noise covariance: 0.1; Depth weighting: Order 0.5 / Maximal amount 10).

To reduce computation time, a lower resolution cortical tessellation (15,000 sources) was used to generate the wMNE source model for the purpose of regional time-frequency analysis. A high-resolution cortical tessellation (150,000 sources) was used for the remainder of the analysis to maximize the spatial resolution.

2.2.5. Time-frequency decomposition

A time-frequency (TF) decomposition was done to select the optimal band-pass filter to apply to the pre-processed IIPT recording before further analysis. This analysis was conducted on all participants using a randomly selected subset consisting of 10% of the presented IITPs. An anatomical ROI was selected for the TF decomposition. Given the putative primary AC's location over HG (Liegeois-Chauvel et al., 1991), the ROI was based on the Desikan-Killiany parcellation for HG generated by FreeSurfer (Desikan et al., 2006), which was then manually enlarged to cover the surrounding sulcal space on both hemispheres. Using the 15,000 source-model, the analyzed ROI overlying HG covered an average of 320.6 sources (SD 23.7) or 49.6 cm^2 (SD 4.28) per participant. Only this step, the TF decomposition, used this lower resolution source model. Once the bandwidth of interest was identified through this step, the following step (Estimation of STRFs), was done on the higher resolution, 150,000 source-model.

The recording was divided into trials of 1 s, from −500 to 500 ms with respect to each IIPT. The DC offset was corrected using the 500 ms period before each IIPT as a baseline. Time-series for each source within the ROI were extracted for each trial. These time-series were then subjected to a TF-decomposition using Morlet wavelets (Tallon-Baudry and Bertrand, 1999) characterized by a central frequency of 1 Hz and a time resolution of 1 s. The decomposition was analyzed in 1 Hz-sized frequency bins. These parameters were chosen to maximize the spectral resolution at the 100 ms response latency (M100), with the goal of using the M100 response for the remainder of the analysis. The M100 is the earliest detectable event-related response attributable to the AC that can be reliably measured with auditory evoked fields (Pantev et al., 1988). Moreover, the M100 response measured by MEG correlates well spatially with the response measured through intracranial recordings (Godey et al., 2001).

The resulting TF-decompositions were then normalized by z-score transformation using a 250 ms-baseline before each IIPT, and an average across all ROI sources for each participant was obtained. A conservative z-score threshold of 1 was applied to the average TF-decomposition to identify the information-containing frequency bands at the 100 ms latency. Across all participants, the minimum lower cutoff frequency was

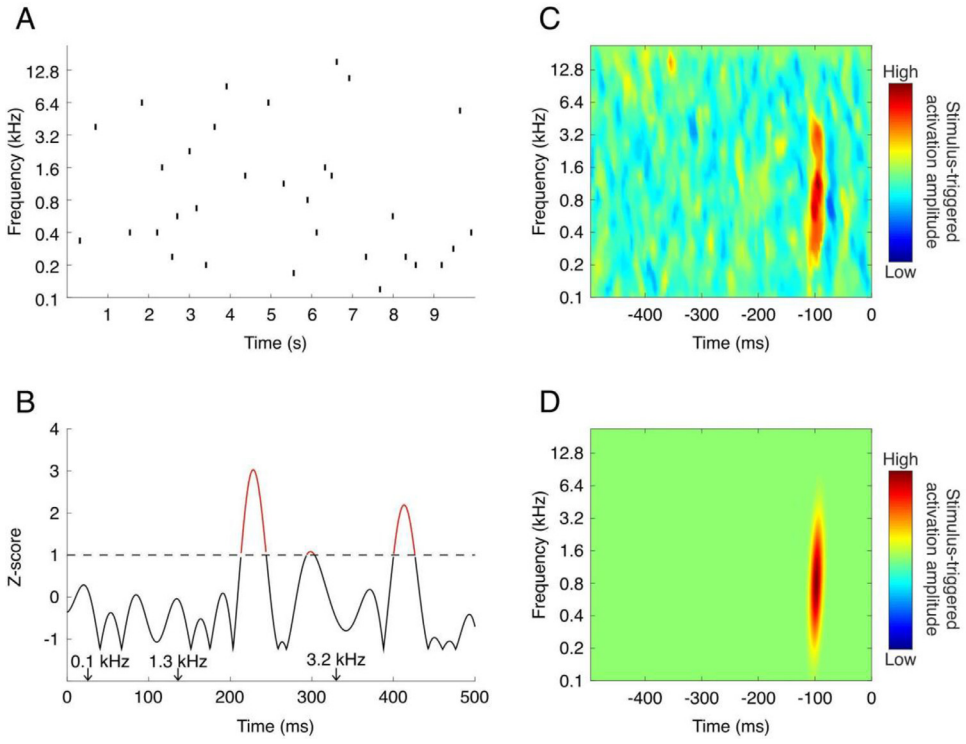

Fig. 2. Generation of STRFs. The process of generating an STRF is depicted. (A) 10-second sample of the IIPT recording. (B) Z-score transformed source-space time series. The threshold for defining a significant activation event is shown with a dashed line at a z-score of 1. Significant activation events are shown in red. The time points at which three sample pure tones were played are shown with arrows above the x-axis. (C) Sample STRF, representing the average stimuli preceding all significant activation events for a given source. This particular STRF has a best frequency of 1.3 kHz. (D) Gaussian-fitted STRF.

3 Hz and the maximum upper cutoff frequency was 13 Hz (for the average TF-decomposition, see Figure S6; for the individual participants' TF-decomposition values, see Table S1). A band-pass filter of 3–13 Hz with a stopband attenuation of 60 dB was therefore applied to the preprocessed IIPT recordings for further analysis.

2.2.6. Estimation of STRFs

The ROI was constructed with the Desikan-Killiany parcellation of the transverse temporal gyrus (HG) and the part of the superior temporal gyrus that is posterior to HG, given that our focus was on the purported region of the primary auditory cortex. STRFs were generated using a technique based on the reverse correlation approach for each source within our ROI. Fig. 2 depicts this process.

The source-space time series was first extracted and converted to absolute values to eliminate the effect of the dipole current's directionality which is not of interest for our application. The absolute value of the time series was then transformed into a z-score normalized time series dynamically by recalculating the mean and standard deviation at each reasonably long segment of silence. This segment of silence had to be at least 100 ms long and begin a minimum of 350 ms after the end of the previous IIPT stimulus to avoid contaminating the baseline with late stimulus-related responses.

From the z-score transformed source-space time-series, local maxima were extracted. A significant activation event was defined as a local-maximum with z-score > 1 (shown in red in Fig. 2, panel B). Such a conservative z-score threshold was chosen to avoid missing activation events that could be reliably time-locked to a stimulus but that may have an amplitude that is relatively low. This choice is counterbalanced by the fact that we weigh activations proportionally to their z-score amplitude, as explained below. We believe this low z-score threshold in combination with a weighting system leads to a more objective selection of significant activations. In comparison, choosing a higher z-score threshold arbitrarily to select a smaller number of activations could be highly dependent on the signal-to-noise ratio of a particular experiment, where different thresholds may lead to different tonotopic maps.

To calculate STRFs, a method based on reverse correlation analysis (deCharms et al., 1998; de Boer and Kuyper 1968) was used. Reverse correlation analysis can be used to reliably estimate a neuron's STRF when the stimulus is uncorrelated, or sampled randomly and uniformly across the spectrotemporal dimensions as is the case with our IIPT stimulus (Theunissen et al., 2000). In summary, the STRF produced through reverse correlation represents the linear estimate of the optimal stimulus preceding a neuronal activation event. It is calculated by computing the average stimulus, in both spectral and temporal dimensions, that precedes a neuronal activation event. For several authors (see for e.g. deCharms et al., 1998), this neuronal activation event is a spike rate, and the STRF quantity is therefore a stimulus-triggered spike rate average. In the method described below, we used a stimulus-triggered activation amplitude average (the average z-score value of the significant activation events), which is more in keeping with the metric being recorded by MEG. The importance of a given activation event on the resulting STRF is therefore proportional to its amplitude.

More specifically, this STRF was computed as a matrix $STRF(f,t)$, where f represents each 32 presented stimulus frequencies and t represents 4 ms bins within the 500 ms time window preceding a significant activation event. For each significant activation event i, the stimulus content in the preceding 500 ms time window was extracted. For each stimulus with frequency f and time t within this time-window, a value corresponding to the z-score amplitude of the corresponding significant activation event was defined as $Z_i(f,t)$. This z-score amplitude was then corrected for the slight variation in the total number of stimuli presented for each stimulus frequency by multiplying it by the coefficient $(f) = \frac{S}{S_f}$, where S represents the mean number of stimuli presented per stimulus frequency, and S_f represents the total number of stimuli presented of frequency f. The corrected z-score activation amplitudes corresponding to each stimulus within the reverse correlation time-windows were then summated to generate the final matrix representing the average stimulus-triggered activation amplitude:

$$STRF(f,t) = \frac{1}{n}\sum_{i=1}^{n}\left[C(f)\cdot Z_i(f,t)\right]$$

The final STRF was smoothed using a gaussian-weighted moving average with a window size of 4 × 4. The M100 response was then de-

fined as the highest spike within a latency window of 80 to 120 ms. The STRF's best frequency was defined as the frequency that elicited the maximal amplitude of activation at a latency corresponding to the M100 response. The best frequency was z-score transformed using the segment of the STRF from −500 to −350 ms for the purpose of determining which STRF showed a significant M100 response (see *Results, Selection of IIPT-Responsive Sources*).

The STRF was finally fitted to a 2D-gaussian surface, aligned on the peak corresponding to the M100 response, in order to smooth the data for estimation of bandwidth and best temporal modulation rate. To normalize its value according to the overall amplitude, the bandwidth was defined as the full spectral width at half maximum of the gaussian-fit and represents the range of frequencies that can elicit an M100 response. The best temporal modulation rate was calculated as $R = (2W)^{-1}$, where W represents the temporal width of the gaussian-fit. The best temporal modulation rate represents a source's preference for a stimulus with a particular temporal modulation.

Only sources with best frequencies ranging from 0.119 kHz to 18.102 kHz were included in the subsequent analysis (total of 30/32 frequencies). The frequency extremes were eliminated to eliminate the edge-effect caused by smoothing the STRFs.

The Brainstorm process used to generate STRFs and map the STRF features onto a cortex surface is available under an open source BSD license at the following GitHub repository: https://github.com/NeuroSensoryBiomarkingLab/MEGACmapping.

3. Results

3.1. Estimation of STRFs

For this analysis, we recorded the neural responses to a 10-minute IIPT stimulus train non-invasively in ten participants (labeled S1 to S10) using a 275-channel whole-head MEG system (CTF MEG International Services Ltd.). MEG records magnetic fields outside the head, and a reverse problem must be solved to localize the source of the magnetic fields from where they originate inside the brain as electrical currents produced by neuronal activity. To do so, we used Weighted Minimum Norm Estimates (wMNE) (Lin et al., 2006), which constrains each source to a one-dimensional perpendicular orientation with respect to a cortex surface obtained through an MRI-based cortical reconstruction generated with FreeSurfer (Dale and Sereno, 1993). Our analysis was conducted on a high cortical tessellation (150,000 sources) to maximize the potential for high spatial resolution.

We generated STRFs for each source within our region of interest (ROI) in the right and left hemispheres of ten participants (S1 to S10) using a technique based on reverse correlation analysis adapted to MEG data and detailed in *Materials and Methods*. The STRF represents the average stimulus-triggered activation amplitude (the average z-score value of every significant neuronal activation event). The resulting STRFs clearly display several important spectrotemporal characteristics expected of neurons in the AC (Fig. 3). These include temporal features such as best temporal modulation rate and response latency, as well as spectral features such as best frequency and frequency bandwidth. The STRFs provide information about the auditory stimuli most likely to elicit a significant response from a given source. While the majority of STRFs had a single peak at a latency of about 100 ms (representing the M100 response), we could identify a number of sources that exhibited a smaller peak at a latency of 50 ms (representing the earlier M50 response). Some sources exhibited complex STRF spectrotemporal patterns, including some with frequency sweeps.

Key properties that can be obtained through analysis of STRFs are shown in Fig. 4. These histograms represent a group-level average among all participants. Best frequencies were represented along a bimodal distribution with one peak at 0.283 kHz and another at 0.8 kHz. However, the range was large, extending throughout all presented frequencies. On average, 90% of sources per participant had a best frequency between 0.2 and 3.2 kHz. Frequency bandwidths were most commonly 2.5 octaves, with the remainder of sources exhibiting a large range of bandwidth. M100 latency was most commonly at 110 ms. Finally, best temporal modulation rates also followed a bimodal distribution, with one peak at 15 Hz (with rates ranging from 10 to 24 Hz), and another centered around 33 Hz (with rates ranging from 25 to 100 Hz). These key properties were extracted from only the right hemisphere, as the identification of tonotopic gradients was more robust in the right cerebral hemisphere of participants (See also Section IV - Investigation of lateralization to IIPT stimuli).

3.2. Selection of IIPT-responsive sources

We defined IIPT-responsive sources as those having an STRF M100 response peak greater than a z-score of 3.5, with a latency between 80 and 120 ms, and a minimum STRF bandwidth of 0.375 octaves (see *Materials and Methods* for precise definitions). The high z-score threshold enables the selection of only those sources that are very IIPT-responsive. This threshold is determined based on the amount of smoothing that is used in the STRF-generation and the signal-to-noise ratio of an experiment. The latency thresholds enable the identification of the M100 response with a range of response latencies. Finally, the bandwidth threshold enables the selection of physiologically plausible receptive fields, eliminating sources that could have a significant "single-bin" receptive field due to chance alone, given the high number of data bins present in the STRF.

3.3. Identification of a tonotopic gradient

To demonstrate the utility of computing STRFs in MEG to study the spatial topographic organization of the auditory cortex, we generated tonotopic maps from the best frequency values of the STRFs for each IIPT-responsive neuronal source. A tonotopic organization could be identified in the right temporal lobe for all participants, as shown in Fig. 5. Because of variability between participants in the position of tonotopic gradient reversals and in the underlying cortical anatomy which covers only a very small area, we do not show a group-level average using currently available tools in the Brainstorm suite, as this leads to loss of valuable gradient information. The gradient pattern is best analyzed individually or, alternatively, using a manual landmark-based averaging method which has proven successful in some fMRI studies (e.g. Humphries et al., 2010).

For the majority of participants (S1 to S8), a primary tonotopic gradient perpendicular to the longitudinal axis of HG could be identified. This primary gradient is most often centered on the posterior part of HG. Among the two participants who did not have a perpendicular gradient progression, S9 had a simple antero-posterior gradient oriented parallel to the longitudinal axis of HG, and S7 had several circular zones of low and high frequencies with a complex organization not observed in other participants. Of note, all participants had a single HG, while S7 had a complete duplication of HG, and S8 had a partial duplication of HG. There was a more variable tonotopic organization present in planum temporale (PT), usually with a relative overrepresentation of low frequencies.

Other characteristics of STRFs can be projected onto the cortical surface, including bandwidth, latency, and temporal modulation. The right-hemisphere maps for these characteristics are presented in Figure S1, S2, and S3. The map stability across time was investigated by comparing best frequency maps in two participants taken 11 days apart. Vertex-wise correlation analyses showed strong map reliability, as evidenced by high determination coefficients (R2) of 0.741 and 0.506 (see also Figure S5).

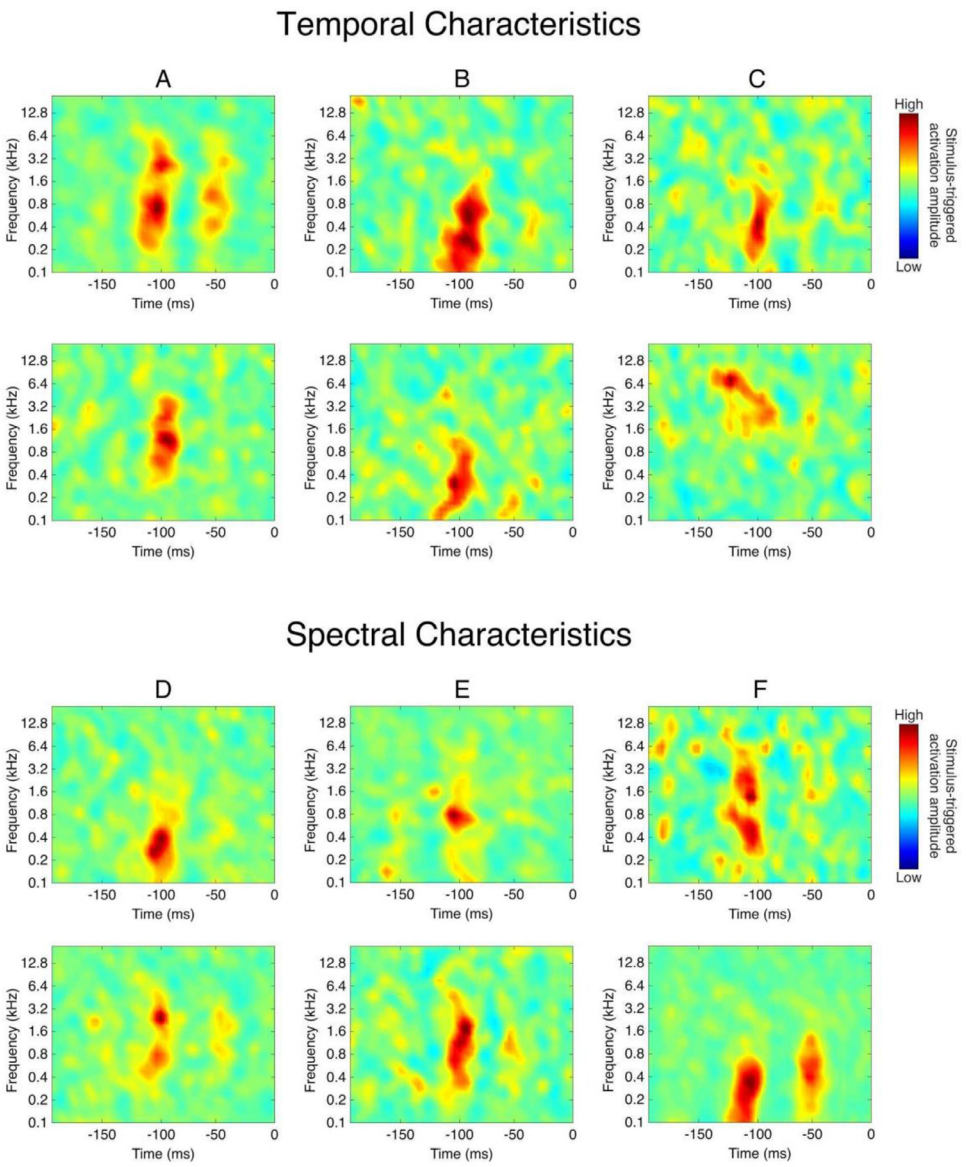

Fig. 3. Variety of STRF characteristics. Sample STRFs exhibiting a range of spectral and temporal characteristics. (A) M50 response (top = strong response; bottom = no response). (B) Best temporal modulation rate (top = low modulation rate; bottom = high modulation rate). (C) Temporal complexity (top = preference for a temporally isolated stimulus; bottom = preference for a downward frequency sweep). (D) Best frequency (top = low frequency; bottom = high frequency). (E) Frequency bandwidth (top = small bandwidth; bottom = large bandwidth). (F) Spectral complexity (top = two spectral peaks eliciting an M100 response; bottom = two spectral peaks eliciting an M100 and/or an M50 response).

3.4. Investigation of lateralization to IIPT stimuli

Identification of tonotopic gradients was more robust in the right cerebral hemisphere of participants, which is why the remainder of our analysis was performed on the right. Best frequency maps for the left hemisphere are shown in Figure S4.

The left hemisphere's tonotopic maps had a decreased signal-to-noise ratio, a shorter range of best frequencies, and less elaborate tonotopic gradients with some participants having no discernible gradient. To investigate whether this was associated with a difference in the number of IIPT-responsive sources in each hemisphere, we calculated the percentage of IIPT-responsive sources within each hemisphere's ROI (Fig. 6). While there were over 50% of IIPT-responsive sources in the left hemisphere's ROI, a two-tailed paired *t*-test revealed that the right hemisphere had 24.3% more IIPT-responsive sources than the left hemisphere (95% CI: 14.8 - 33.8; $p = 0.0003$), confirming a lateralization to the right.

4. Discussion

We have described a novel method for functional mapping of the human AC using MEG, showing that it can reliably extract important information about the spatial organization of auditory processing when using STRFs generated from a dense pure tone auditory stimulus.

4.1. Estimation of STRFs using MEG

MEG has been used in the past to successfully generate physiologically plausible STRFs using discrete (Constantino et al., 2017) and continuous stimuli (Crosse et al., 2016; Ding and Simon, 2012), but to the best of our knowledge, this is the first *in vivo* MEG study to estimate STRFs using reverse correlation in human participants for the purpose of mapping the functional organization of the AC. Our findings support that it is possible to generate physiologically plausible STRFs with excellent variety in terms of spectrotemporal patterns, akin to what has been reported in other mammalian studies (see for e.g.: Elhilali et al., 2007; Massoudi et al., 2015). Some STRFs exhibited complex spectrotemporal patterns, which is a testament to the high temporal resolution and sufficient spatial resolution of MEG for isolating a variety of neuronal sub-populations in a relatively small cortical region. In some cases, we could even detect a strong M50 response.

Our analysis of the distribution of these STRF properties revealed a bimodal distribution of best frequency centered on 0.283 kHz and 0.8 kHz, with the vast majority of sources (90%) having a best fre-

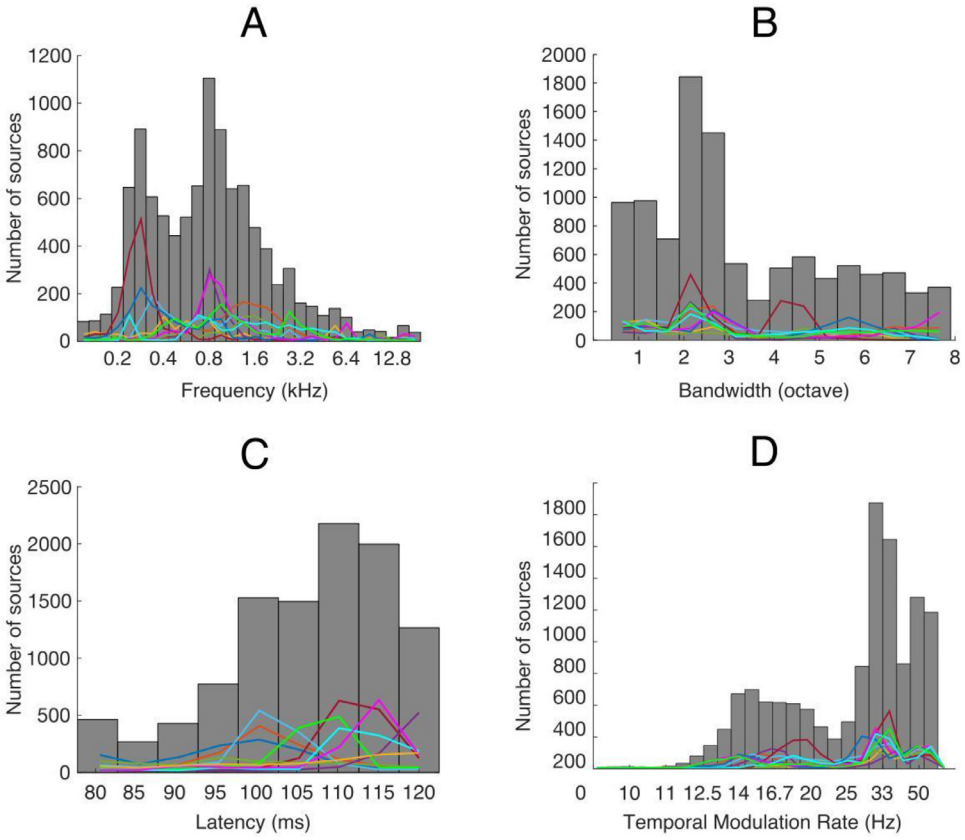

Fig. 4. Histograms of STRF characteristics. Histograms showing the total number of sources from all 10 participants for each of the following STRF characteristics: best frequency (A), bandwidth (B), latency (C), and best temporal modulation rate (D). Lines representing the individual contribution of each participant are superimposed onto the histograms.

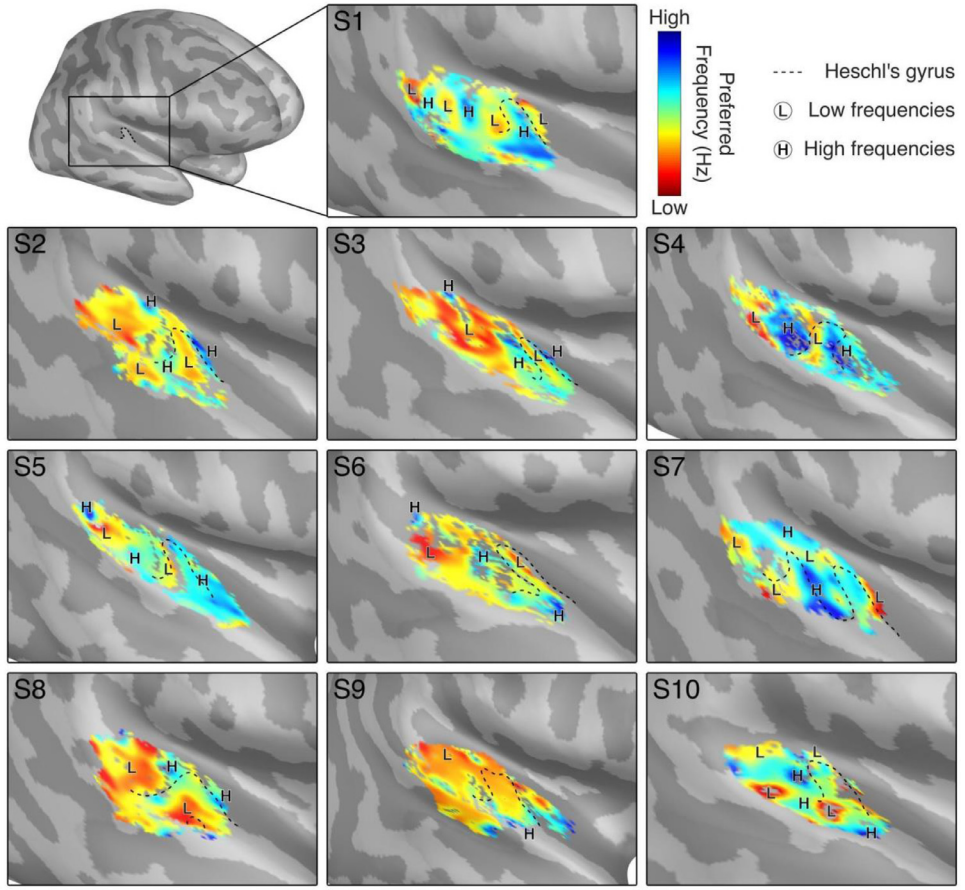

Fig. 5. Best frequency maps and tonotopic gradient organization. Best frequency maps are shown for participants 1 to 10. Major regions of high and low frequencies are marked as H and L, respectively, and Heschl's gyrus is outlined for reference. Tonotopic gradients can be identified in all participants. Colormap limits are set near the local minima and maxima of each participant to best visualize gradient patterns.

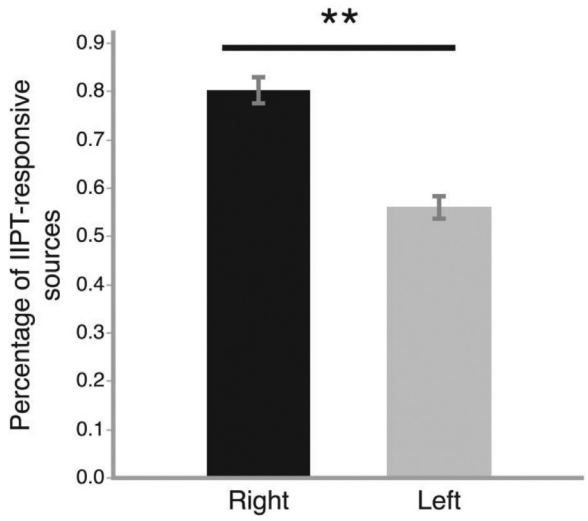

Fig. 6. Average percentage of IIPT-responsive sources for each cerebral hemisphere among all participants. The average percentage of IIPT-responsive sources is presented for the right and left hemispheres. The percentage represents the number of IIPT-responsive sources over all possible sources within our region of interest. The error bars represent the standard error to the mean. Two-tailed paired *t*-test results in $p = 0.0003$, t-value −5.7887, DF = 9.

quency between 0.2 kHz and 3.2 kHz. Our results are comparable with those of a small study of four epilepsy patients with intracranial electrode recordings, where only the frequencies 0.32–3.2 kHz elicited neuronal responses, and 0.25–2.0 kHz elicited the strongest responses (Bitterman et al., 2008). This finding likely reflects emphasis on frequencies used in speech sounds. In an articulation test (intelligibility of speech communication), participants scored 95% accuracy when a low-pass filter of 4.0 kHz was applied to speech sounds and 100% when the low-pass filter was 7.0 kHz (Monson et al., 2014), suggesting that spectral content below 7.0 KHz is the most important for speech comprehension. The wider audible frequency range in humans extends up to about 20 kHz for young healthy individuals (Monson et al., 2014), and while these were represented in our dataset, they were markedly underemphasized when compared to the frequency band of speech. The E-A-RTONE 3A earphones are a potential confounder, given that its frequency responses, although audible, progressively decrease in intensity beyond 3 kHz. While this could contribute to the underrepresentation of higher frequencies in our dataset, the fact that frequency representation begins dropping well below 3 kHz (starting at 0.8 kHz) suggests that the underrepresentation is truly representative of the underlying functional organization.

The ability to extract STRFs using MEG is significant. The STRFs we produced are in keeping with what is physiologically expected of neurons in the human AC, and their characteristics are likely important in understanding the subdivisions of neuronal populations in humans. STRF analyses have also been crucial to better understand the mechanisms underlying plasticity. In mice, STRFs have provided an excellent means of visualizing the changes in spectrotemporal characteristics of neuronal response over time (Kamal et al., 2013). Until now, correlating STRFs with spatial organization in the auditory cortex was only possible in animal studies and intracranial recording studies in humans, but our proposed methodology provides a novel non-invasive method to do so in humans.

4.2. MEG-generated tonotopic maps

A mirror-symmetric tonotopic gradient has been described in most fMRI studies (Da Costa et al., 2011; Formisano et al., 2003; Humphries et al., 2010; Langers and van Dijk, 2012; Moerel et al., 2012), and the majority of our participants exhibit a very similar gradient pattern, usually centered around a region of low frequency in the posterior part of HG. The directionality of the gradient found using our method is most closely aligned with the findings of Humphries et al. (2010), Da Costa et al. (2011), Formisano et al. (2003), Moerel et al. (2012), who also describe a primary gradient perpendicular to the longitudinal axis of HG (though Langers and van Dijk (2012) in contrast describe a latero-medial progression) . Furthermore, a review of fMRI and cyto- and myeloarchitectural studies proposed a model of the human AC with a tonotopic gradient oriented at a similar angle with respect to HG (Moerel et al., 2014). Therefore, the similarities between the tonotopic organization we describe and that found in the fMRI literature supports the accuracy of our technique. It also adds to the body of evidence pointing to a primary gradient centered on HG that is perpendicularly oriented to the longitudinal axis of HG, likely representing the primary AC.

The IIPT stimulus has notable advantages over more standard and sparser presentation rates used in fMRI and MEG literature (Lütkenhöner et al., 2003; Cha et al., 2016). Because of its temporally dense nature, it allows for the presentation of a significantly larger amount of pure tones in a short time frame, which in turn allows for a larger number of different frequencies to be presented, enabling the production of a more finely graduated tonotopic map. For example, a typical auditory cortex mapping study will use 8 different frequencies and require 30 min of tone presentations (Cha et al., 2016). In contrast, the IIPT can fit up to 32 different frequencies presented within a 10-minute sequence. As a result, we believe that the IIPT stimulus contributes to map quality by achieving a greater spectral resolution than other methods within a given data acquisition time frame.

This supports the hypothesis that the spatial resolution of MEG is sufficient to study tonotopic gradients in the human AC, and allows us to leverage MEG's excellent temporal resolution to study short-latency-dependent events and complex spectrotemporal characteristics inherent in auditory processing that are impossible to study using fMRI. The method of choice to investigate the spatial organization of the auditory cortex has more often than not been fMRI (Su et al., 2014). Prior EEG and MEG studies investigating this topic generally used a limited tone diversity or didn't capture the sources' characteristics beyond the preferred frequency (Romani et al., 1982; Pantev et al., 1996; Cansino et al., 2003). Although more recent studies have investigated the temporal responses by also representing STRFs via specialized algorithms such as boost and spatiotemporal searchlight representational similarity analysis, they have generally focused on speech-related stimuli and/or do not attempt to map information extracted from the STRF back on the cortex to study tonotopy (David et al., 2007; Su et al., 2014; Constantino et al., 2017). A significant advantage of our current method is that the short IIPT train enables a finer stimulus spectral resolution combined with an increased number of repetitions for each presented frequency, thus producing finer and more robust tonotopic maps. Furthermore, the use of STRFs further reduces the stimulus presentations required for other receptive field representations such as the tuning curve. To achieve similar results with other source analysis methods, data acquisition time would need to be significantly increased.

One drawback of this method is that data analysis is computationally expensive, due to the need to compute an STRF for every vertex in the ROI.

There was significant inter-participant variability in our dataset, which is consistent with the findings of methods boasting greater spatial resolution such as fMRI (Humphries et al., 2010). Nonetheless, we could still identify consistent tonotopic gradient progressions that shared similar patterns and directionality among the majority of participants. These patterns extend from the core auditory cortex to the putative location of the belt and parabelt areas. These other subfields have already been characterized based on tonotopy (Moerel et al., 2012), but the technique presented here has the spatial resolution that would allow further tem-

poral characterization of these subfields by harnessing the MEG's high temporal resolution.

4.3. Right-hemispheric lateralization of response to pure tones

The tonotopic maps produced using our methodology have led us to identify a right-hemispheric lateralization of the tonotopic organization in response to IIPTs at M100. Despite there being over 50% of sources in the left hemisphere that responded to the IIPT stimulus, the characterization of a tonotopic organization was less robust than in the right hemisphere, with some participants having no discernible gradient. While functional lateralization of the human AC has been extensively studied with respect to stimuli involving music and speech sounds (Tervaniemi and Hugdahl, 2003), lateralization of tonotopy using pure tone stimuli has received less attention in the literature. In a single fMRI study, the presence of a clearer tonotopic organization was noted in the right primary AC compared to the left, although there was significant inter-participant variability (Langers et al., 2007). Right-sided specialization for frequency-specific tuning has also been noted in intracranial recordings of auditory evoked potentials (Liégeois-Chauvel et al., 2001) and in a previous study using MEG (Ozaki and Hashimoto, 2007). There is also evidence pointing to left-ear advantage (and therefore right hemispheric lateralization) when human participants are presented with tonal, but not noise stimuli (Sininger and Bhatara, 2012).

The bulk of the evidence on pitch and music points to the right hemisphere having better spectral resolution (Zatorre et al., 2002), and therefore implicating it more in music, pitch and tonal processing. This contrasts with the left hemisphere's better temporal resolution, rendering it more important in the processing of much faster temporal variations in the sound amplitude envelope, as is the case in speech. These hypotheses are supported by lesioning studies showing that lesions affecting the right HG result in deficits in the perception of pitch, by electrophysiological studies showing an association between pitch perception and the timing of cortical activity in the right hemisphere, and by a variety of functional imaging studies showing a predilection for tonal processing in the right hemisphere (Zatorre et al., 1994, 1992; Perry et al., 1999; Halpern and Zatorre, 1999; Griffiths et al., 1999; Penhune et al., 1998; Hugdahl et al., 1999; Tervaniemi et al., 2000).

Several fMRI studies have identified a tonotopic organization in the left hemisphere (see for e.g. Formisano et al., 2003; Talavage et al., 2004; Langers et al., 2007). While we could identify a tonotopic organization in a subset of participants' left hemispheres, this was less robust than on the right. This discrepancy could be due to at least two reasons. First, it is possible that the type of stimulus could be implicated. We used a spectrotemporally dense pure tone stimulus with a much higher presentation rate than is typically used in fMRI studies. However, because the left hemisphere is thought to be important in the processing of temporal characteristics of sound (Zatorre et al., 2002), it would be difficult to explain why such a difference in the stimulus presentation rate could result in a lateralization to the right hemisphere. Second, the discrepancy could be related to the timing of acquisition and the temporal resolution of the two modalities. With MEG, the high temporal resolution allows us to isolate specific auditory cortical responses such as the M100 response, whereas the BOLD response used in fMRI results from neuronal activity occurring over a much longer time period, dictated by hemodynamic properties. Therefore, the activity captured through fMRI may relate to activity taking place much later than the M100 response in the auditory processing hierarchy. We believe this to be the more likely explanation behind this observation.

4.4. Limitations

There are limitations to the method we propose. First, the sound intensity (volume) of stimulus presentation is a limiting factor in the ability to resolve a tonotopic gradient. In order to truly capture the characteristic (best) frequency of a neuron, the lowest sound intensity that will elicit a response must be found; however, current electrophysiological and functional neuroimaging techniques are not sensitive enough to record neuronal responses barely above threshold, and therefore require the use of higher sound intensities. Coupled with the notion that neurons respond to a broader range of frequencies when stimulated by higher sound intensities (Recanzone, 2000), doing so may result in the spread of activation limiting the accuracy and resolvability of the measured tonotopic gradient (Tanji et al., 2010). While this limitation cannot be avoided, we used A-weighted stimulus intensity to compensate for the differences in volume necessary to lead to equivalent intensity perception at each frequency (Fletcher and Munson, 1933).

There are possible artifacts related to recording auditory evoked fields in the region of the AC. MEG is selectively sensitive to current along the walls of sulci, and cannot detect current at the crest of gyri and bottom of sulci (Puce and Hämäläinen, 2017). Moreover, the activity recorded from regions lying in close proximity to other surfaces, as is the case with the AC, could potentially be altered or even canceled by conflicting currents occurring simultaneously on the adjacent surface (Ahlfors et al., 2010). In our dataset, we did not observe any deficiency in the identification of IIPT-responsive sources in the crests of gyri and bottom of sulci. This leads us to believe that any potential alteration in signal occurring as a consequence of the macroanatomy of the AC did not prevent adequate source estimation with MEG.

Although our analysis is based on the earliest consistently detectable response in MEG (Pantev et al., 1988), the M100 response, what it represents remains controversial. Intracranial recordings have localized M100 to the lateral portion of HG and PT (Godey et al., 2001; Liégeois-Chauvel et al., 1994), while non-invasive recordings have localized it exclusively to PT (Lütkenhöner and Steinsträter, 1998; Engelien et al., 2000), which may be interpreted as activity originating from secondary ACs. This evidence rightfully has led some to question the claim that M100 originates from the primary AC (Moerel et al., 2014). However, our data is not entirely consistent with this view. We show that there is clear activity at M100 along the purported anatomical location of the primary AC, HG. There are two possible ways to reconcile these differences. It may be that the spatial resolution of MEG is such that activity in spatially separated cortical areas appears to be overlapping. In this case, most of the observed activity could be originating from PT but falsely appear to be extending beyond PT into HG. We believe this is unlikely, particularly given that tonotopic gradients were identified as progressing in shorter distance increments than the distance between PT and HG. Another possibility is that primary and secondary auditory processing are overlapping in some regions of the AC. If this were the case, it would indicate that HG is both involved in primary and secondary processing. There is evidence showing that the earlier 50 ms-latency response (M50) is in fact located within the same anatomical region as the M100 response (Wang et al., 2014), which could support this hypothesis. Even if M100 represents higher order processing, we assume, as others have (Su et al., 2014), that the tonotopic organization of auditory processing should in theory remain stable over at least several hundred milliseconds. Even if it does not, investigation of the M100 response using MEG remains valuable, as insights into later auditory processing steps can be gained from studying responses in secondary ACs.

4.5. Conclusions

Here, we show that MEG can be used to characterize the tonotopic organization of the AC by generating STRFs with a spectrotemporally dense pure tone stimulus. We described a large variety of STRF patterns consistent with the expected variety of neuronal subtypes that can be further studied both spectrally, through measures such as frequency bandwidth, and temporally, through measures such as best temporal modulation rate, and latency. The best frequency maps and tonotopic gradients we were able to generate shared strong similarities with those observed in other fMRI studies. MEG therefore is able to provide sufficient spatial resolution to study the spatial functional organization of the

human AC, including the microarchitecture of auditory subfields, while providing additional benefits through its high temporal resolution. Our proposed method has significant implications for the field of auditory processing, as it is the first to effectively capture both high spatial resolution and spectrotemporal information, which together provide a more complete understanding of auditory processing in humans.

Declaration of Competing Interest

none

Credit authorship contribution statement

Jean-Pierre R. Falet: Methodology, Software, Visualization, Writing – original draft. **Jonathan Côté:** Conceptualization, Data curation, Formal analysis, Investigation, Project administration, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Veronica Tarka:** Data curation, Formal analysis, Investigation, Writing – review & editing. **Zaida Escila Martínez-Moreno:** Data curation, Formal analysis, Investigation. **Patrice Voss:** Writing – review & editing. **Etienne de Villers-Sidani:** Conceptualization, Funding acquisition, Supervision.

Acknowledgments

We thank Sylvain Baillet, PhD, Robert Zatorre, PhD, and Kuwook Cha, PhD, from the Department of Neurology and Neurosurgery at McGill, for providing helpful comments about our methods and analysis.

Data availability statement

A sample of our dataset is freely available for download from the OpenNeuro platform at the following link: https://openneuro.org/datasets/ds003082/versions/1.0.0. All other data can be provided upon request by contacting the corresponding author, Jonathan Côté, by email.

Funding statement

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) [grant number RGPIN-2019–04761]; the Centre for Research on Brain, Language and Music (CRBLM); and the Réseau québécois de recherche sur le vieillissement (RQRV). The funding sources were not involved in the study.

Ethics approval statement

This study was approved by the research ethics board of the Montreal Neurological Institute.

Patient approval statement

All participants provided written informed consent.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118222.

References

Abrams, D.A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D.J., Menon, V., 2011. Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. Cereb. Cortex 21 (7), 1507–1518. doi:10.1093/cercor/bhq198.

Aguirre, G.K., Zarahn, E., D'Esposito, M., 1998. The variability of human, BOLD hemodynamic responses. Neuroimage 8 (4), 360–369.

Ahlfors, S.P., Han, J., Lin, F.-.H., Witzel, T., Belliveau, J.W., Hämäläinen, M.S., Halgren, E., 2010. Cancellation of EEG and MEG signals generated by extended and distributed sources. Hum. Brain Mapp. 31 (1), 140–149. doi:10.1002/hbm.20851.

Alain, C., Woods, D.L., Covarrubias, D., 1997. Activation of duration-sensitive auditory cortical fields in humans. Electroencephalogr. Clin. Neurophysiol. Evoked Potentials Sect. 104 (6), 531–539.

Barth, D.S., Sutherling, W., Broffman, J., Beatty, J., 1986. Magnetic localization of a dipolar current source implanted in a sphere and a human cranium. Electroencephalogr. Clin. Neurophysiol. 63 (3), 260–273. doi:10.1016/0013-4694(86)90094-5.

Beukes, E.W., Munro, K.J., Purdy, S.C., 2009. Duration-sensitive neurons in the auditory cortex. Neuroreport 20 (13), 1129–1133.

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 (7175), 197–201. doi:10.1038/nature06476.

Blake, D.T., Merzenich, M.M., 2002. Changes of AI receptive fields with sound density. J. Neurophysiol. 88 (6), 3409–3420. doi:10.1152/jn.00233.2002.

Calabrese, A., Schumacher, J.W., Schneider, D.M., Paninski, L., Woolley, S.M.N., 2011. A generalized linear model for estimating spectrotemporal receptive fields from responses to natural sounds. PLoS One 6 (1), e16104. doi:10.1371/journal.pone.0016104.

Cansino, S., Ducorps, A., Ragot, R., 2003. Tonotopic cortical representation of periodic complex sounds. Hum. Brain Mapp. 20 (2), 71–81. doi:10.1002/hbm.10132.

Carlin, M.A., Elhilali, M., 2015. Modeling attention-driven plasticity in auditory cortical receptive fields. Front. Comput. Neurosci. 9, 106. doi:10.3389/fncom.2015.00106.

Cha, K., Zatorre, R.J., Schönwiesner, M., 2016. Frequency selectivity of voxel-by-voxel functional connectivity in human auditory cortex. Cereb. Cortex 26 (1), 211–224. doi:10.1093/cercor/bhu193.

Constantino, F.C., Villafañe-Delgado, M., Camenga, E., Dombrowski, K., Walsh, B., & Simon, J.Z. (2017, July 27). Functional significance of spectrotemporal response functions obtained using magnetoencephalography. doi:10.1101/168997

Cote, J., de Villers-Sidani, E., 2020. Auditory cortex mapping dataset. OpenNeuro. https://openneuro.org/datasets/ds003082/versions/1.0.0. (2020).

Crosse, M.J., Di Liberto, G.M., Bednar, A., Lalor, E.C., 2016. The multivariate temporal response function (mTRF) toolbox: a MATLAB toolbox for relating neural signals to continuous stimuli. Front. Hum. Neurosci. 10, 604. doi:10.3389/fnhum.2016.00604.

Da Costa, S., van der Zwaag, W., Marques, J.P., Frackowiak, R.S.J., Clarke, S., Saenz, M., 2011. Human primary auditory cortex follows the shape of Heschl's gyrus. J. Neurosci. 31 (40), 14067–14075. doi:10.1523/JNEUROSCI.2000-11.2011.

Dale, A.M., Sereno, M.I., 1993. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. J. Cognit. Neurosci. 5 (2), 162–176.

David, S., Mesgarani, N., Shamma, S., 2007. Estimating sparse spectro-temporal receptive fields with natural stimuli. Network 18m, 191–212. doi:10.1080/09548980701609235.

de Boer, E., Kuyper, P., 1968. Triggered correlation. IEEE Trans. Biomed. Eng. BME-15 (3), 169–179. doi:10.1109/tbme.1968.4502561.

deCharms, R.C., Blake, D.T., Merzenich, M.M., 1998. Optimizing sound features for cortical neurons. Science 280 (5368), 1439–1443. doi:10.1126/science.280.5368.1439.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., … Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 1;31 (3), 968–980. doi:10.1016/j.neuroimage.2006.01.021.

de Villers-Sidani, E., Alzghoul, L., Zhou, X., Simpson, K.L., Lin, R.C., Merzenich, M.M., 2010. Recovery of functional and structural age-related changes in the rat primary auditory cortex with operant training. Proc. Natl. Acad. Sci. 107 (31), 13900–13905. doi:10.1073/pnas.1007885107, http://dx.doi.org/.

Ding, N., Simon, J.Z., 2012. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. J. Neurophysiol. 107 (1), 78–89. doi:10.1152/jn.00297.2011.

Dobri, S.G.J., Ross, B., 2021. Total GABA level in human auditory cortex is associated with speech-in-noise understanding in older age. Neuroimage 225, 117474. doi:10.1016/j.neuroimage.2020.117474.

Elhilali, M., Fritz, J.B., Chi, T.S., Shamma, S.A., 2007. Auditory cortical receptive fields: Stable entities with plastic abilities. J. Neurosci. 27 (39), 10372–10382. doi:10.1523/JNEUROSCI.1462-07.

Engelien, A., Schulz, M., Ross, B., Arolt, V., Pantev, C., 2000. A combined functional in vivo measure for primary and secondary auditory cortices. Hear. Res. 148 (1–2), 153–160. doi:10.1016/s0378-5955(00)00148-9.

Franco, A., Eberlen, J., Destrebecqz, A., Cleeremans, A., Bertels, J., 2015. Rapid serial auditory presentation: a new measure of statistical learning in speech segmentation. J. Exp. Psychol. 62 (5), 346–351. doi:10.1027/1618-3169/a000295.

Fischl, B. (2013). Freesurfer homepage. Retrieved 4 January 2020, from http://surfer.nmr.mgh.harvard.edu/

Fischl, B., Liu, A., Dale, A.M., 2001. Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. IEEE Trans. Med. Imaging 20 (1), 70–80. doi:10.1109/42.906426.

Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis. Neuroimage doi:10.1006/nimg.1998.0396.

Fletcher, H., Munson, W.A., 1933. Loudness, its definition, measurement and calculation. Bell Syst. Tech. J. doi:10.1002/j.1538-7305.1933.tb00403.x.

Formisano, E., Kim, D.S., Di Salle, F., van de Moortele, P.F., Ugurbil, K., Goebel, R., 2003. Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron 40 (4), 859–869. doi:10.1016/s0896-6273(03)00669-x.

Gardumi, A., Ivanov, D., Havlicek, M., Formisano, E., Uludağ, K., 2017. Tonotopic maps in human auditory cortex using arterial spin labeling. Hum. Brain Mapp. 38 (3), 1140–1154. doi:10.1002/hbm.23444.

Godey, B., Schwartz, D., de Graaf, J.B., Chauvel, P., Liégeois-Chauvel, C., 2001. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. 112 (10), 1850–1859. doi:10.1016/s1388-2457(01)00636-8.

Griffiths, T.D., Johnsrude, I., Dean, J.L., Green, G.G., 1999. A common neural substrate for the analysis of pitch and duration pattern in segmented sound? Neuroreport 10 (18), 3825–3830. doi:10.1097/00001756-199912160-00019.

Halpern, A.R., Zatorre, R.J., 1999. When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. Cereb. Cortex 9 (7), 697–704. doi:10.1093/cercor/9.7.697.

Huang, C.W., Huang, M.X., Ji, Z., Swan, A.R., Angeles, A.M., Song, T., Huang, J.W., Lee, R.R., 2016. High-resolution MEG source imaging approach to accurately localize Broca's area in patients with brain tumor or epilepsy. Clin. Neurophysiol. 127 (5), 2308–2316. doi:10.1016/j.clinph.2016.02.007.

Huang, X.M., Dale, A.M., Song, T., Halgren, E., Harrington, D.L., Podgorny, I., Canive, J.M., Lewis, S., Lee, R.R., 2006. Vector-based spatial–temporal minimum L1-norm solution for MEG. Neuroimage 31 (3), 1025–1037. doi:10.1016/j.neuroimage.2006.01.029.

Huang, M.X., Mosher, J.C., Leahy, R.M., 1999. A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. Phys. Med. Biol. 44 (2), 423–440. doi:10.1088/0031-9155/44/2/010.

Hugdahl, K., Brønnick, K., Kyllingsbaek, S., Law, I., Gade, A., Paulson, O.B., 1999. Brain activation during dichotic presentations of consonant-vowel and musical instrument stimuli: a 15O-PET study. Neuropsychologia 37 (4), 431–440. doi:10.1016/s0028-3932(98)00101-8.

Humphries, C., Liebenthal, E., Binder, J.R., 2010. Tonotopic organization of human auditory cortex. Neuroimage 50 (3), 1202–1211. doi:10.1016/j.neuroimage.2010.01.046.

Huotilainen, M., Tiitinen, H., Lavikainen, J., Ilmoniemi, R.J., Pekkonen, E., Sinkkonen, J., … Näätänen, R., 1995. Sustained fields of tones and glides reflect tonotopy of the auditory cortex. Neuroreport 6 (6), 841–844. doi:10.1097/00001756-199504190-00004.

Joachimsthaler, B., Uhlmann, M., Miller, F., Ehret, G., Kurt, S., 2014. Quantitative analysis of neuronal response properties in primary and higher-order auditory cortical fields of awake house mice (M us musculus). Eur. J. Neurosci. 39 (6), 904–918.

Kamal, B., Holman, C., de Villers-Sidani, E., 2013. Shaping the aging brain: role of auditory input patterns in the emergence of auditory cortical impairments. Front. Syst. Neurosci. 7, 52. doi:10.3389/fnsys.2013.00052.

Kowalski, N., Depireux, D.A., Shamma, S.A., 1996. Analysis of dynamic spectra in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary dynamic spectra. J. Neurophysiol. 76 (5), 3524–3534. doi:10.1152/jn.1996.76.5.3524.

Kuriki, S., Murase, M., 1989. Neuromagnetic study of the auditory responses in right and left hemispheres of the human brain evoked by pure tones and speech sounds. Exp. Brain Res. Exp. Hirnforsch. Exp. Cereb. 77 (1), 127–134. doi:10.1007/bf00250574.

Langers, D.R.M., Backes, W.H., van Dijk, P, 2007. Representation of lateralization and tonotopy in primary versus secondary human auditory cortex. Neuroimage 34 (1), 264–273. doi:10.1016/j.neuroimage.2006.09.002.

Langers, D.R.M., van Dijk, P., 2012. Mapping the tonotopic organization in human auditory cortex with minimally salient acoustic stimulation. Cereb. Cortex 22 (9), 2024–2038. doi:10.1093/cercor/bhr282.

Leahy, R.M., Mosher, J.C., Spencer, M.E., Huang, M.X., Lewine, J.D., 1998. A study of dipole localization accuracy for MEG and EEG using a human skull phantom. Electroencephalogr. Clin. Neurophysiol. 107 (2), 159–173. doi:10.1016/S0013-4694(98)00057-1.

Leaver, A.M., Rauschecker, J.P., 2016. Functional topography of human auditory cortex. J. Neurosci. 36 (4), 1416–1428. doi:10.1523/jneurosci.0226-15.2016.

Liégeois-Chauvel, C., Giraud, K., Badier, J.M., Marquis, P., Chauvel, P., 2001. Intracerebral evoked potentials in pitch perception reveal a functional asymmetry of the human auditory cortex. Ann. N. Y. Acad. Sci. 930, 117–132. doi:10.1111/j.1749-6632.2001.tb05728.x.

Liégeois-Chauvel, C., Musolino, A., Badier, J.M., Marquis, P., Chauvel, P., 1994. Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. Electroencephalogr. Clin. Neurophysiol. 92 (3), 204–214. doi:10.1016/0168-5597(94)90064-7.

Liegeois-Chauvel, C., Musolino, A., Chauvel, P., 1991. Localization of the primary auditory area in man. Brain J. Neurol. 114 (Pt 1A), 139–151. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1900211 .

Linden, J.F., Liu, R.C., Sahani, M., Schreiner, C.E., Merzenich, M.M., 2003. Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. J. Neurophysiol. 90 (4), 2660–2675. doi:10.1152/jn.00751.2002.

Lin, F.-H., Witzel, T., Ahlfors, S.P., Stufflebeam, S.M., Belliveau, J.W., Hämäläinen, M.S., 2006. Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. Neuroimage 31 (1), 160–171. doi:10.1016/j.neuroimage.2005.11.054.

Lütkenhöner, B., Krumbholz, K., Seither-Preisler, A., 2003. Studies of tonotopy based on wave N100 of the auditory evoked field are problematic. Neuroimage 19 (3), 935–949. doi:10.1016/S1053-8119(03)00172-1.

Lütkenhöner, B., Steinsträter, O., 1998. High-precision neuromagnetic study of the functional organization of the human auditory cortex. Audiol. Neurootol. 3 (2–3), 191–213. doi:10.1159/000013790.

Massoudi, R., Van Wanrooij, M.M., Versnel, H., Van Opstal, A.J., 2015. Spectrotemporal response properties of core auditory cortex neurons in awake monkey. PLoS One 10 (2), e0116118. doi:10.1371/journal.pone.0116118.

Moerel, M., De Martino, F., Formisano, E., 2012. Processing of natural sounds in human auditory cortex: tonotopy, spectral tuning, and relation to voice sensitivity. J. Neurosci. 32 (41), 14205–14216. doi:10.1523/JNEUROSCI.1388-12.2012.

Moerel, M., De Martino, F., Formisano, E., 2014. An anatomical and functional topography of human auditory cortical areas. Front. Neurosci. 8, 225. doi:10.3389/fnins.2014.00225.

Monson, B.B., Hunter, E.J., Lotto, A.J., Story, B.H., 2014. The perceptual significance of high-frequency energy in the human voice. Front. Psychol. 5, 587. doi:10.3389/fpsyg.2014.00587.

Nagel, K.I., Doupe, A.J., 2008. Organizing principles of spectro-temporal encoding in the avian primary auditory area field L. Neuron 58 (6), 938–955. doi:10.1016/j.neuron.2008.04.028.

Nasiotis, K., Clavagnier, S., Baillet, S., Pack, C.C., 2017. High-resolution retinotopic maps estimated with magnetoencephalography. Neuroimage 145 (Pt A), 107–117. doi:10.1016/j.neuroimage.2016.10.017.

Niranjan, A., Laing, E.J.C., Laghari, F.J., Richardson, R.M., Lunsford, L., D., 2013. Preoperative magnetoencephalographic sensory cortex mapping. Stereotact. Funct. Neurosurg. 91, 314–322. doi:10.1159/000350019.

Okada, Y.C., Lahteenmäki, A., Xu, C., 1999. Experimental analysis of distortion of magnetoencephalography signals by the skull. Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. 110 (2), 230–238. doi:10.1016/s0013-4694(98)00099-6.

Ozaki, I., Hashimoto, I., 2007. Human tonotopic maps and their rapid task-related changes studied by magnetic source imaging. Can. J. Neurol. Sci. 34 (2), 146–153. doi:10.1017/s0317167100005965.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., Elbert, T., 1995. Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. Electroencephalogr. Clin. Neurophysiol. 94 (1), 26–40. doi:10.1016/0013-4694(94)00209-4.

Pantev, C., Hoke, M., Lehnertz, K., Lütkenhöner, B., Anogianakis, G., Wittkowski, W., 1988. Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. Electroencephalogr. Clin. Neurophysiol. 69 (2), 160–170. doi:10.1016/0013-4694(88)90211-8.

Pantev, C., Hoke, M., Lehnertz, K., Lütkenhöner, B., Fahrendorf, G., Stöber, U., 1990. Identification of sources of brain neuronal activity with high spatiotemporal resolution through combination of neuromagnetic source localization (NMSL) and magnetic resonance imaging (MRI). Electroencephalogr. Clin. Neurophysiol. 75, 173–184. doi:10.1016/0013-4694(90)90171-F.

Pantev, C., Roberts, L.E., Elbert, T., Roβ, B., Wienbruch, C., 1996. Tonotopic organization of the sources of human auditory steady-state responses. Hear. Res. 101 (1–2), 62–74. doi:10.1016/S0378-5955(96)00133-5.

Penhune, V.B., Zattore, R.J., Evans, A.C., 1998. Cerebellar contributions to motor timing: a PET study of auditory and visual rhythm reproduction. J. Cognit. Neurosci. 10 (6), 752–765. doi:10.1162/089892998563149.

Perry, D.W., Zatorre, R.J., Petrides, M., Alivisatos, B., Meyer, E., Evans, A.C., 1999. Localization of cerebral activity during simple singing. Neuroreport 10 (18), 3979–3984. doi:10.1097/00001756-199912160-00046.

Puce, A., Hämäläinen, M.S., 2017. A review of issues related to data acquisition and analysis in EEG/MEG studies. Brain Sci. 7 (6). doi:10.3390/brainsci7060058.

Recanzone, G.H., 2000. Spatial processing in the auditory cortex of the macaque monkey. Proc. Natl. Acad. Sci. U.S.A. 97 (22), 11829–11835. doi:10.1073/pnas.97.22.11829.

Regan, D., 1989. Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine. Elsevier. Retrieved from https://books.google.com/books/about/Human_brain_electrophysiology.html?hl=&id=5dVqAAAAMAAJ.

Romani, G.L., Williamson, S.J., Kaufman, L., 1982. Tonotopic organization of the human auditory cortex. Science 216 (4552), 1339–1340. doi:10.1126/science.7079770.

Ross, B., Dobri, S., Schumann, A., 2020. Speech-in-noise understanding in older age: the role of inhibitory cortical responses. Eur. J. Neurosci. 51, 891–908. doi:10.1111/ejn.14573.

Schreiner, C.E., Polley, D.B., 2014. Auditory map plasticity: diversity in causes and consequences. Curr. Opin. Neurobiol. 24 (1), 143–156. doi:10.1016/j.conb.2013.11.009.

Sen, K., Theunissen, F.E., Doupe, A.J., 2001. Feature analysis of natural sounds in the songbird auditory forebrain. J. Neurophysiol. 86 (3), 1445–1458. doi:10.1152/jn.2001.86.3.1445.

Sininger, Y.S., Bhatara, A., 2012. Laterality of basic auditory perception. Laterality 17 (2), 129–149. doi:10.1080/1357650X.2010.541464.

Su, L., Zulfiqar, I., Jamshed, F., Fonteneau, E., Marslen-Wilson, W., 2014. Mapping tonotopic organization in human temporal cortex: representational similarity analysis in EMEG source space. Front. Neurosci. 8, 368. doi:10.3389/fnins.2014.00368.

Tadel, F. (2019, December 23). Introduction. Retrieved 4 January 2020, from https://neuroimage.usc.edu/brainstorm/

Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. Comput. Intell. Neurosci. 2011, 879716. doi:10.1155/2011/879716.

Talavage, T.M., Edmister, W.B., 2004. Nonlinearity of FMRI responses in human auditory cortex. Hum. Brain Mapp. 22 (3), 216–228. doi:10.1002/hbm.20029.

Talavage, T.M., Sereno, M.I., Melcher, J.R., Ledden, P.J., Rosen, B.R., Dale, A.M., 2004. Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. J. Neurophysiol. 91 (3), 1282–1296. doi:10.1152/jn.01125.2002.

Tallon-Baudry, C., Bertrand, O., 1999. Oscillatory gamma activity in humans and its role in object representation. Trends Cognit. Sci. 3 (4), 151–162. doi:10.1016/s1364-6613(99)01299-1.

Tanji, K., Leopold, D.A., Ye, F.Q., Zhu, C., Malloy, M., Saunders, R.C., Mishkin, M., 2010. Effect of sound intensity on tonotopic fMRI maps in the unanesthetized monkey. Neuroimage 49 (1), 150–157. doi:10.1016/j.neuroimage.2009.07.029.

Tervaniemi, M., Hugdahl, K., 2003. Lateralization of auditory-cortex functions. Brain Res. Brain Res. Rev. 43 (3), 231–246. doi:10.1016/j.brainresrev.2003.08.004.

Tervaniemi, M., Medvedev, S.V., Alho, K., Pakhomov, S.V., Roudas, M.S., Van Zuijen, T.L., Näätänen, R., 2000. Lateralized automatic auditory processing of phonetic versus musical information: a PET study. Hum. Brain Mapp. 10 (2), 74–79. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10864231 .

Tesche, C.D., Uusitalo, M.A., Ilmoniemi, R.J., Huotilainen, M., Kajola, M., Salonen, O., 1995. Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. Electroencephalogr. Clin. Neurophysiol. 95 (3), 189–200. doi:10.1016/0013-4694(95)00064-6.

Theunissen, F.E., Sen, K., Doupe, A.J., 2000. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. J. Neurosci. 20 (6), 2315–2331. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10704507 .

Uusitalo, M.A., Ilmoniemi, R.J., 1997. Signal-space projection method for separating MEG or EEG into components. Med. Biol. Eng. Comput. 35 (2), 135–140. doi:10.1007/bf02534144.

van Wassenhove, V., Nagarajan, S.S., 2007. Auditory cortical plasticity in learning to discriminate modulation rate. J. Neurosci. 27 (10), 2663–2672. doi:10.1523/JNEUROSCI.4844-06.2007.

Verkindt, C., Bertrand, O., Perrin, F., Echallier, J.F., Pernier, J., 1995. Tonotopic organization of the human auditory cortex: N100 topography and multiple dipole model analysis. Electroencephalogr. Clin. Neurophysiol. 96 (2), 143–156. doi:10.1016/0168-5597(94)00242-7.

Wang, Y., Feng, Y., Jia, Y., Wang, W., Xie, Y., Guan, Y., … Huang, L., 2014. Auditory M50 and M100 sensory gating deficits in bipolar disorder: a MEG study. J. Affect. Disord. 152-154, 131–138. doi:10.1016/j.jad.2013.08.010.

Weisz, N., Wienbruch, C., Hoffmeister, S., Elbert, T., 2004. Tonotopic organization of the human auditory cortex probed with frequency-modulated tones. Hear. Res. 191 (1–2), 49–58. doi:10.1016/j.heares.2004.01.012.

Woods, D.L., Herron, T.J., Cate, A.D., Yund, E.W., Stecker, G.C., Rinne, T., Kang, X., 2010. Functional properties of human auditory cortical fields. Front. Syst. Neurosci. 4, 155. doi:10.3389/fnsys.2010.00155.

Woolley, S.M.N., Gill, P.R., Theunissen, F.E., 2006. Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. J. Neurosci. 26 (9), 2499–2512. doi:10.1523/JNEUROSCI.3731-05.2006.

Zatorre, R.J., Belin, P., Penhune, V.B., 2002. Structure and function of auditory cortex: music and speech. Trends Cognit. Sci. 6 (1), 37–46. doi:10.1016/s1364-6613(00)01816-7.

Zatorre, R.J., Evans, A.C., Meyer, E., 1994. Neural mechanisms underlying melodic perception and memory for pitch. J. Neurosci. 14 (4), 1908–1919. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8158246 .

Zatorre, R.J., Evans, A.C., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch discrimination in speech processing. Science 256 (5058), 846–849. doi:10.1126/science.1589767.