NeuroImage 279 (2023) 120330

Contents lists available at ScienceDirect

NeuroImage

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

A high-density EEG and structural MRI source analysis of the frequency following response to missing fundamental stimuli reveals subcortical and cortical activation to low and high frequency stimuli

Karl D. Lerud [a,*], Roeland Hancock [b], Erika Skoe [c]

[a] University of Maryland College Park, Institute for Systems Research, 20742, United States of America
[b] Yale University, Wu Tsai Institute, 06510, United States of America
[c] University of Connecticut, Department of Speech, Language, and Hearing Sciences, Cognitive Sciences Program, 06269, United States of America

ARTICLE INFO

Dataset link: https://osf.io/2ar9q/

Keywords:
Frequency following response
auditory
EEG
source analysis
subcortical
brainstem
missing fundamental
envelope
auditory cortex
MNE
minimum norm
Brainstorm

ABSTRACT

Pitch is a perceptual rather than physical phenomenon, important for spoken language use, musical communication, and other aspects of everyday life. Auditory stimuli can be designed to probe the relationship between perception and physiological responses to pitch-evoking stimuli. One technique for measuring physiological responses to pitch-evoking stimuli is the frequency following response (FFR). The FFR is an electroencephalographic (EEG) response to periodic auditory stimuli. The FFR contains nonlinearities not present in the stimuli, including correlates of the amplitude envelope of the stimulus; however, these nonlinearities remain undercharacterized. The FFR is a composite response reflecting multiple neural and peripheral generators, and their contributions to the scalp-recorded FFR vary in ill-understood ways depending on the electrode montage, stimulus, and imaging technique. The FFR is typically assumed to be generated in the auditory brainstem; there is also evidence both for and against a cortical contribution to the FFR. Here a methodology is used to examine the FFR correlates of pitch and the generators of the FFR to stimuli with different pitches. Stimuli were designed to tease apart biological correlates of pitch and amplitude envelope. FFRs were recorded with 256-electrode EEG nets, in contrast to a typical FFR setup which only contains a single active electrode. Structural MRI scans were obtained for each participant to co-register with the electrode locations and constrain a source localization algorithm. The results of this localization shed light on the generating mechanisms of the FFR, including providing evidence for both cortical and subcortical auditory sources.

1. Introduction

The frequency following response is a scalp-recorded electrical potential originating in the auditory periphery and brain that both mirrors and adds to auditory stimuli in many ways, eliciting multiple frequency response components. It is a well-characterized and useful tool for studying auditory function in both research and clinical settings (Skoe and Kraus, 2010). However a number of its aspects remain elusive or underexplained. For instance, although the generators of the FFR have long been a topic of interest (Sohmer et al., 1977; Gardi et al., 1979; Coffey et al., 2016, 2017; Farahani et al., 2017; Bidelman, 2018; Coffey et al., 2019; Teichert et al., 2022), there is still little consensus about the peripheral and neural generators of the FFR. It is also not well known which of these generators lead to which FFR frequency components (e.g., even- vs. odd-order components, Gockel et al., 2011; Gnanateja and Maruthy, 2019). A novel experimental paradigm was thus conceived and applied to address these gaps in the literature. Using carefully-controlled missing fundamental-type stimuli, in combination with high-density EEG FFR source analysis, this study sought to ascertain the subcortical and cortical generators of the FFR at different stimulus fundamental frequencies, and also to tease apart the nonlinear even- from odd-order components by presenting stimuli in alternating polarity. The even-order component of the FFR is often associated with the response envelope of a complex, multi-component stimulus, while the odd-order component is often associated with the stimulus fine structure.

Much of the existing research on the FFR generators has assumed that this response is fully generated in subcortical auditory structures (cochlear nucleus (CN), superior olive (SOC), and inferior colliculus (IC)), and ventured to determine the relative contribution of each structure in scalp recordings. Sohmer et al. (1977) utilized clinical

* Corresponding author.
E-mail address: lerud@umd.edu (K.D. Lerud).

https://doi.org/10.1016/j.neuroimage.2023.120330
Received 2 April 2023; Received in revised form 29 July 2023; Accepted 14 August 2023
Available online 19 August 2023

populations with lesions or hearing loss and determined the main source of the FFR to be the IC, a similar conclusion reached in invasive animal studies such as Smith et al. (1975). On the other hand, Gardi et al. (1979) used invasive approaches in cats and concluded that the dominant source was the CN, with lesser contributions from the IC and the cochlear microphonic (a peripheral component).

There have also been more recent efforts utilizing modern neuroimaging techniques. Bidelman (2015) used a 64-channel EEG paradigm and determined the main source of the FFR was the IC, although this work did not benefit from individual head models, which may improve source localization accuracy. However Coffey et al. (2016) performed a high-density MEG study utilizing individual anatomical MRI scans (T1s) to localize the FFR to a speech syllable sound (100 Hz fundamental frequency (f0)) and found a bilateral cortical contribution from Heschl's gyrus, in addition to sources in the auditory brainstem and thalamus (medial geniculate body). Gorina-Careta et al. (2021) used a concurrent MEG/EEG experimental paradigm and came to similar conclusions regarding the cortical and subcortical sources through the MEG source analysis (EEG source analysis was not done in that study). Compared to EEG, MEG is more sensitive to superficial sources and tangential source dipoles in general, while EEG is sensitive to both radial and tangential sources and is comparatively more sensitive to deep sources (Ahlfors et al., 2010; Goldenholz et al., 2008). Thus, it is possible that MEG emphasizes cortical source results more than EEG. Additionally, because these results contradict previously-held assumptions about the FFR being exclusively a subcortical response, it is important to attempt to replicate them using EEG and stimulus f0s that span widely in frequency. Therefore, an experiment was undertaken to be the first high-density EEG FFR source analysis study using low and high f0s (80 Hz and 210 Hz). For each f0, two different stimuli were constructed, providing a means for comparing the source estimates across stimuli. This study is meant in part to replicate the results of Coffey et al. (2016) and Gorina-Careta et al. (2021), but also to attempt to shed new light on the FFR by precisely characterizing the sources of the nonlinear responses to the f0.

2. Methods

2.1. Overview

This is the first FFR study to attempt source localization with EEG and individual anatomical data; to accomplish this, the study's data acquisition first involved a 1 mm T1-weighted anatomical MRI scan. Next, participants were fitted with a 256-electrode net manufactured by EGI (Eugene, OR), using an EEG amplifier capable of a high sampling rate appropriate for FFRs. Before EEG acquisition, twelve pictures were taken of each participant's head with the net firmly fitted using EGI's hardware and software. These pictures aid in localizing the electrodes along with three fiducial points for co-registration with the MRI scan. EEG data then comprised roughly 60 GB per participant, thus processing was done on a computing cluster specialized for high-memory computational needs. Once EEG data was downsampled and trials averaged, the remainder of the analysis and source localization was done on personal computers.

2.2. Participants

After a single pilot participant was run to test the functionality of the acquisition and data analysis pipelines, twelve participants were recruited for this experiment. Participants were recruited based on the fact that all had previously undergone FFR acquisition in unrelated studies and were known to have a robust FFR. Having a relatively low signal-to-noise ratio in an FFR to complex sounds does not indicate any problem of processing or function, nor does having a higher signal-to-noise ratio necessarily indicate greater perceptual abilities. However for this study's purposes, better FFRs were preferred to more accurately characterize the various responses, as desired. The average age of the participants was 23 ($SD = 2.76$) and all were female. All participants had normal hearing thresholds ($\leq$ 20 dB HL across octave frequencies from 250 Hz to 8 kHz) as had been previously established from audiometric measures when they participated in earlier FFR studies. Participants were monetarily compensated, including the pilot participant. Informed consent was obtained from all participants and the study design was approved by the University of Connecticut Institutional Review Board.

2.3. General study design

These participants had never been exposed to these particular stimuli before. All data acquisition including EEG and MRI for each subject was done in one 2-hour session as efficiency of experimental structure was important. All data was acquired at the University of Connecticut Brain Imaging Research Center (BIRC). Upon arriving, participants first read consent documents and signed them, and also filled out a brief questionnaire about music and language background. They also filled out an MRI safety screening form. Immediately upon finishing these documents, their head circumference was measured so the appropriate EGI EEG acquisition net could begin soaking in a potassium chloride electrolytic solution while the participant's MRI was being acquired. After head measurement, participants were led to the MRI room where the BIRC MRI technician checked their MRI safety screening form and performed the MRI acquisition.

Participants were then led to an EEG acquisition room. Outside the sound proof booth, the soaked EEG net was put on their head and fastened. They were then led to another room that housed the EGI photogrammetry hardware and software, and twelve pictures were acquired from different angles to co-register the EEG electrode locations with the MRI data for source analysis. After photograph acquisition, EEG acquisition took place. Binaural ear inserts were placed in participants' ears to deliver the stimuli and they began watching a silent movie. No behavior of any kind was elicited or measured, other than a request to minimize eye blinks and body movements during perception of the stimuli. After stimulus delivery and EEG acquisition began, EEG acquisition was approximately 51 min, including 20-second rest periods between each of eight blocks of stimulus delivery. Participants were then debriefed and compensated.

2.4. Stimuli

FFRs were recorded to four shifted missing fundamental (MF) stimuli, two with a missing f0 of 80 Hz and the other two with an f0 of 210 Hz. 80 Hz and 210 Hz were selected because previous work suggested that both f0s produce robust FFRs but that the relative contribution of the cortex would be less for 210 Hz compared to 80 Hz (Tichko and Skoe, 2017). Each stimulus was a 350-millisecond (ms) three-component complex with linear on and off ramps of 5 ms and each consisted of three primary frequencies (f1, f2, f3) that were nominal harmonics 2, 3, and 4 of an MF, shifted up by an irrational ratio (Table 1). When played together these primaries produce a response at the envelope frequency, which is mathematically the difference between successive primaries (f2-f1, f3-f2), and is more technically known as the quadratic difference tone (QDT). To be able to separate the response to the QDT from the response to the primaries, the primaries were shifted up by an irrational amount. Shifted MF stimuli were selected because Gockel et al. (2011) found that they generated a strong even-order FFR at the MF frequency. Simply described, an MF stimulus is one in which harmonics of a fundamental frequency are physically present, but the fundamental itself is not present. A shifted MF stimulus is one in which the primaries have all been shifted by a constant frequency, while retaining the frequency interval between harmonics. The amount of frequency shift in such stimuli can be expressed as a

Table 1
A table specifying the parameters of the four stimuli used in the present study. Frequencies are in units of Hertz, and the irrational numbers are approximate. Stimulus primaries are harmonics 2, 3, and 4 of a missing fundamental specified as the QDT, shifted up by the specified ratio. For all these stimuli, the envelope frequency is equal to the QDT.

	Primary 1	Primary 2	Primary 3	Shift ratio	QDT	CDT
Stimulus 1	178.8562	258.8562	338.8562	$\frac{\sqrt{2}}{6}$	80	98.8562
Stimulus 2	174.1421	254.1421	334.1421	$\frac{\sqrt{2}}{8}$	80	94.1421
Stimulus 3	469.4975	679.4975	889.4975	$\frac{\sqrt{2}}{6}$	210	259.4975
Stimulus 4	457.1231	667.1231	877.1231	$\frac{\sqrt{2}}{8}$	210	247.1231

ratio of the frequency change to the original fundamental. The stimuli in Gockel et al. (2011) were shifted by ratios of $\frac{1}{2}$ and $\frac{1}{4}$, but to be sure which frequencies in the FFR were related to the "envelope" of the stimulus, the present study utilized irrational shift ratios (Table 1).

Even- and odd-order portions of the FFR are obtained by delivering stimuli in two opposite polarities, averaging those two groups of trials separately, and then summing the two polarity conditions to obtain the even-order response, and subtracting them to obtain the odd-order response. Historically, many FFR experiments have utilized the alternating polarity technique to avoid electromagnetic stimulus contamination in the electrodes (i.e., "stimulus artifact") (Chimento and Schreiner, 1990). The even-order portion of the FFR is often associated with the envelope of the stimulus because it indeed often contains a prominent component at the amplitude envelope frequency of the stimulus, as well as some harmonics. In a harmonic stimulus such as a speech syllable, the envelope frequency is also the frequency difference between successive harmonics, so even if there is little or no energy at the fundamental frequency in the stimulus, the auditory system produces prominent energy there. It is in fact these nonlinear relationships between stimulus components, such as difference or summation, that predicts their presence in the brain's response (Lerud et al., 2014), rather than just the amplitude envelope of the stimulus.

Four different pitch-shifted stimuli were used, summarized in Table 1. Some expected nonlinearities are specified in the Table as well, including the quadratic difference tone (QDT), which is also sometimes referred to as the missing fundamental frequency. The usage of the word "tone" for such nonlinear auditory phenomena is rooted in the recording of mechanical and pressure waves produced by the auditory periphery (Bian and Chen, 2008; Zwicker, 1979; Nuttall and Dolan, 1993), but the terminology has persisted, referring to the same frequencies, within studies recording electromagnetic auditory responses such as EEG (Bhagat and Champlin, 2004; Gockel et al., 2012). The perceived pitch of such shifted stimuli is not the same as the QDT frequency (Schouten et al., 1962; Moore and Moore, 2003), but perception was not studied here. The QDT, which is in general the difference frequency between successive stimulus frequencies, will be the primary response of interest in the present study. For these stimuli, the QDT response coincides with the missing f0 of the stimulus. Because a response is generated at a frequency that is not in the stimulus, it is classified as a nonlinear response. For the purposes of the current work, "QDT response" and "f0 response" will be used synonymously.

Each trial was 350 ms long, followed by a random inter-stimulus interval with a mean of 137.5 ms and a possible spread of 25 ms in both the negative and positive directions. Inter-stimulus intervals came from a uniform distribution. The four stimuli were delivered in both polarities, thus the experiment consisted of eight unique stimuli. These eight stimuli were presented 750 times each, in a pseudorandom order distributed throughout eight blocks. Thus each block was approximately six minutes, and there was also a 20-second break between each block. The random parameters in the stimulus presentation were intended to minimize any habituation that might occur due to short-term plasticity, thus theoretically maximizing the FFR's signal-to-noise ratio. The stimuli were delivered in the same order for each participant.

The stimuli were generated as `.wav` files in MATLAB (The Mathworks Inc., MA, USA). The right channels of the files contained the stimuli and were split and delivered to both ears, while the left channel contained a DC trigger event marker that was delivered along with the EEG data to ensure timing accuracy. The stimuli were delivered with Etymotic ER-2 transducers and 1.5-foot-long tube phones with insert earphones and foam tips. The transducers were mu-metal-shielded to minimize stimulus artifact contamination, and hung approximately 1.5 ft below the EEG electrodes along the participant's torso. The stimuli were calibrated with a physical sound level meter to be 70 dB SPL at the eardrum, and were presented using an external Focusrite USB sound card.

2.5. EEG and MRI acquisition and processing

EEG data was amplified with an MR-compatible, high-impedance EGI NetAmps 410 amplifier. The purpose of using this amplifier was not to collect data in-scanner, but to achieve a higher output sampling rate than the more typical NetAmps 400. Thus EEG data was recorded at the 410's maximum rate of 20,000 Hz, a much more desirable rate for FFR acquisition than the maximum rate of 1000 Hz for the NetAmps 400. All NetAmps amplifiers utilize a hardware lowpass anti-aliasing filter at the Nyquist frequency prior to analog-to-digital conversion. EEG was recorded using EGI's Net Station software. Participants watched a silent movie on a screen approximately one meter in front of their heads while EEG data was being collected. They were instructed to minimize eye and body movements while they were hearing the auditory stimuli, but were otherwise not instructed to behave or react to anything. Once the EEG net was fitted to their head, they were taken to a room with an EGI Geodesic Photogrammetry System (GPS) and pictures were taken for electrode localization with EGI's GPS Solver software. An example of what the fitted EGI EEG net looks like within the solver software is provided in Figs. 1A and 1B. Before acquisition, the net preparation was finished by ensuring that all electrode sponges were in as close contact as possible with the scalp. More electrolytic solution was individually applied to any sponges that required it. Electrode impedances were kept below 50 kΩ, which is typical for a high-impedance system such as EGI's.

EEG processing was done in MATLAB using the free and open source (FOSS) EEG/MEG analysis software package FieldTrip (Oostenveld et al., 2011). At the initial sampling rate of 20,000 Hz, each block of EEG was approximately 15 GB in memory when loaded in MATLAB, which is prohibitive on a normal computer. Thus UConn's High Performance Computing (HPC) cluster was utilized for EEG processing in FieldTrip. Upon initial data analysis, an electrical artifact at 1000 Hz and its harmonics was discovered which contaminated all blocks in all subjects. This electrical contamination was subsequently identified as polling from one or more USB lines or ports from the stimulus delivery computer outside the sound proof EEG booth, likely introduced through the transducer cables. Because brain data beyond this frequency could not be easily analyzed if present, and because the initial rate of 20,000 Hz is unwieldy and greater than necessary, subsequent analysis resampled all EEG to 2500 Hz. Resampling was done with FieldTrip, which in turn utilizes MATLAB's `resample()` function, which uses a lowpass anti-aliasing filter at the new Nyquist frequency prior to downsampling. This sampling rate yields a Nyquist frequency of 1250 Hz, which is sufficient for the present FFR analysis

purposes, and is much greater than the 500 Hz Nyquist frequency of the EGI NetAmps 400 amplifier (sampling rate of 1000 Hz), which would not be sufficient for the present purposes. After resampling, data was subjected to a frequency-domain bandstop filter at the electrical supply AC frequency of 60 Hz and its harmonics. The main filter was a bandpass between 63 Hz and 950 Hz. The low cutoff was arrived at through visual inspection of the filter's magnitude response to be the highest possible frequency that also does not affect the lowest frequency expected in the EEG responses, namely the lower QDT of 80 Hz. The high cutoff was the highest frequency that also completely filtered out the electrical artifact at 1000 Hz. A linear-phase, 500th-order, finite impulse response filter was implemented as zero-phase with MATLAB's `filtfilt()` function and utilized. This filter was applied to the continuous EEG data before epoching, and also had the effect of making all the data zero-mean, which is expected from any sufficiently aggressive highpass filter.

Also before epoching, an automated channel repair procedure was applied to each block of the EEG data. A FASTER-like (Nolan et al., 2010) algorithm selected statistically-outlying channels that needed repair, and FieldTrip's own channel repair function was then utilized. This function replaces the data in each of the outlying channels with a spline interpolation of the data from neighboring channels. The median number of channels repaired per participant was $Mdn = 6.5$, with median absolute deviation $MAD = 1.5$.

The continuous EEG data was then epoched according to the event times that were recorded. There were eight blocks for each subject consisting of 750 trials each. There were four different stimuli, each with two polarity conditions, and these eight unique stimuli were distributed pseudorandomly throughout the eight blocks. After epoching, all EEG channels were re-referenced to a global average for the purpose of later source analysis. Two methods of trial rejection were then applied. The first was another FASTER-like algorithm to detect statistically outlying trials. The second rejected trials based on a threshold of absolute amplitude, under the assumption that trials with very high amplitude were likely contaminated with eye or muscle artifact. The threshold for the latter routine was 65 μV. The number of rejected trials for both of these methods were different; the FASTER method ($Mdn = 13$, $MAD = 4$) typically rejected slightly more than the absolute amplitude method ($Mdn = 11$, $MAD = 9$). Trials selected for rejection based on either or both methods were all rejected from further analysis. Once artifact rejection was complete, trials were then grouped according to which stimulus they corresponded to, and ERPs for each stimulus were generated by averaging over trials. For each stimulus, the even-order portion of the response was then created as the sum of the responses to the two polarity conditions, and the odd-order portion was created as the difference between them.

In order to construct accurate and individualized forward models for source analysis, each participant underwent an anatomical T1-weighted magnetization prepared rapid gradient MRI scan before EEG collection (TR = 2300 ms; TE = 2.98 ms; 1 mm^2 isotropic resolution). The MRIs were collected on a Siemens Prisma 3 T scanner using a 64-channel Head/Neck coil. The FOSS package FreeSurfer (Fischl et al., 2002; Fischl, 2012) was used to segment the entire brain from the raw MRI image. Neocortex was reconstructed as a surface of vertices, each with an orientation, and subcortical structures were also parcellated as a volume of vertices without orientations. The results of FreeSurfer's reconstructions were then imported into the FOSS brain imaging package Brainstorm (Tadel et al., 2011), along with the ERP matrices for each stimulus, for source analysis.

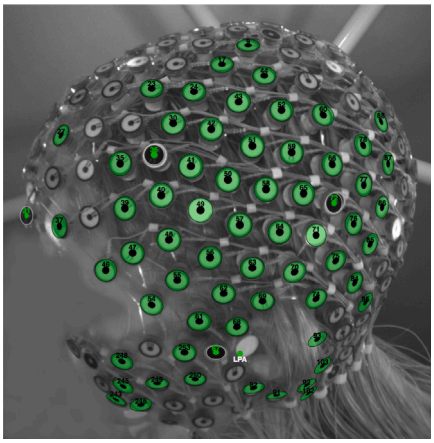
2.6. EEG source analysis of the FFR

After import into Brainstorm, an affine transformation matrix to the MNI152 template brain was created for each FreeSurfer reconstruction. A mixed source model, which contains both surface (orientation-constrained) and volume (not orientation-constrained) vertices, was created in Brainstorm for each subject's head. The entire brain except the cerebellum was used for each source model, containing approximately 24,000 vertices per brain. A forward model was then calculated for each subject using the Boundary Element Method as implemented in the FOSS package OpenMEEG (Gramfort et al., 2010, 2011) which works within Brainstorm. A forward model is a transformation that gives a scalp-space representation of time series data given the source-space data. Typically, however, researchers want a model that does the opposite: Given the scalp-space time series data, a transformation that yields the source-space data is desired. This approximation is called an inverse model, and the forward model is required as a first step to obtaining it.

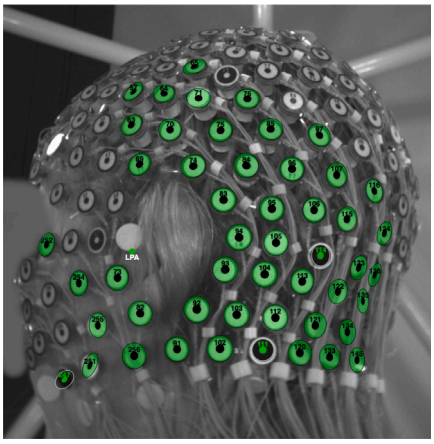
For the source computation with an inverse model to be meaningful, several regions of interest (ROIs) within the source model were established. An ROI, or scout in Brainstorm, consists of several vertices whose time series will be averaged once an inverse model is applied. The scouts were chosen here to be substantially similar to those of both Coffey et al. (2016) and Gorina-Careta et al. (2021), in an attempt to replicate findings indicating that while subcortical auditory structures are the primary producer of the FFR, there is also a non-negligible cortical contribution to the FFR from primary auditory cortex. For this analysis, non-auditory areas served as control cortical regions, with the prediction that energy at the QDT would be near zero in these areas. In total, four cortical scouts and three subcortical scouts were selected. The cortical scouts were: separate left and right primary auditory cortices, combined bilateral frontal poles, and combined bilateral occipital poles. The subcortical were: combined bilateral cochlear nuclei (CN), combined bilateral inferior colliculi (IC), and combined bilateral medial geniculate bodies (MGB). In this context, "combined" again means averaged, just as with the time series of the individual vertices.

The frontal pole scout was taken from the Desikan-Killiany atlas (Desikan et al., 2006), as were both primary auditory scouts, labeled by identifying the transverse temporal (Heschl's) gyrus in both hemispheres. The occipital pole scout was created manually. Cortical scouts created for source analysis are depicted in one subject's brain in Figs. 1E to 1H. None of the atlases parcel out the brainstem or thalamus beyond identifying them as two different structures, and thus subcortical scouts were also created manually by the first author. The CN was created by noting the caudal base of the pons, the IC by noting the prominent anatomical features of the corpora quadrigemina, and the MGB by noting the caudal, medial portions of the bilateral thalamus. All seven scouts were made to be exactly 50 vertices by pruning or growing the scout through a nearest-neighbor search. Subcortical scouts created for source analysis are depicted in one subject's brain in Figs. 1I to 1L. Auditory and frontal scouts were all identified automatically by FreeSurfer for each participant, so there were no transformations necessary for those scouts. The occipital pole scout was drawn manually in one subject and transformed to the rest of the subjects through each of their MNI brain transformation matrices. Subcortical scouts were entered by specifying the seed vertices for each one with MNI coordinates that were then transformed to each specific brain, and grown to 50 vertices.

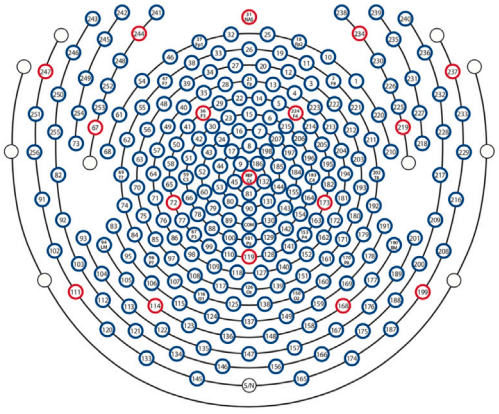
Once the scouts were chosen for source analysis, an inverse model was created for each subject to transform the scalp-space data to time series in source space. A minimum norm estimate (MNE) was used to accomplish this (Gramfort et al., 2014) utilizing the dynamical Statistical Parametric Mapping (dSPM) measure (Dale et al., 2000) made available in Brainstorm's options for minimum norm imaging. The dSPM measure requires an approximation of a noise covariance matrix of all electrodes; this was obtained by concatenating all inter-trial, non-stimulus-related periods of EEG and computing a covariance matrix from this for each subject.



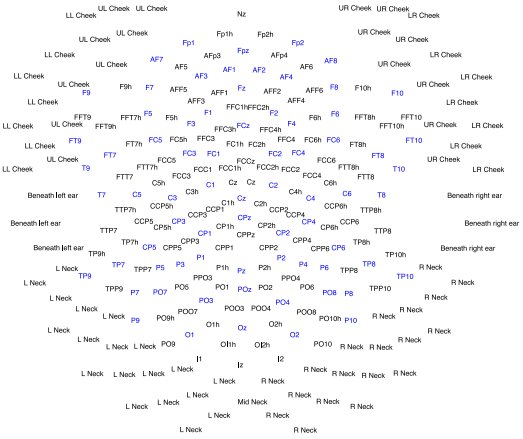
(**A**) View from camera 6.



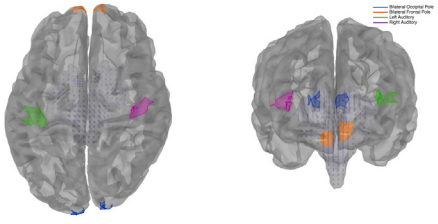
(**B**) View from camera 10.



(**C**) EGI's sensor numbers and locations.

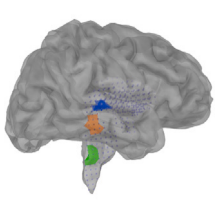


(**D**) Corresponding 10/5 labels and locations.

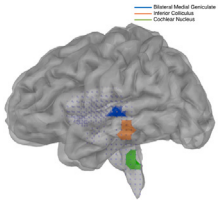


(**E**) Cortical scouts from above.

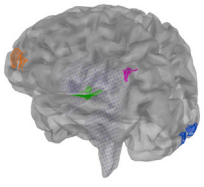
(**F**) View from front.



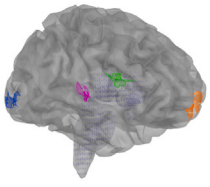
(**I**) Subcortical scouts from right.



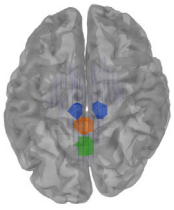
(**J**) View from left.



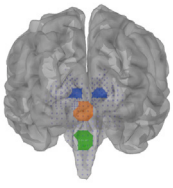
(**G**) View from left.



(**H**) View from right.



(**K**) View from top.



(**L**) View from back.

Fig. 1. A–B: Two views of the fitted EGI Geodesic net from EGI GPS Solver electrode localization software. C–D: A novel 10/5 system of sensor location labels for a geodesic sensor net. Sensors in blue in D were already labeled in EGI's documentation; all others were labeled manually by the authors. E-L: Multiple views of one participant's brain of all cortical and subcortical ROIs.

When the inverse model is calculated, a kernel is created to transform scalp-space EEG data to any desired vertices in the source model. Thus the source-space time series for each scout were calculated and analyzed for their frequency content. Each scout contained 50 vertices, so a single time series was obtained from each scout by averaging the time series of the individual vertices that comprise each scout. The volume (subcortical) vertices do not have an orientation constraint, so Brainstorm gives three time series for each of those instead of one. Each of these represents an orthogonal axis of orientation. Frequency content of volume scouts was thus calculated by averaging each of the three

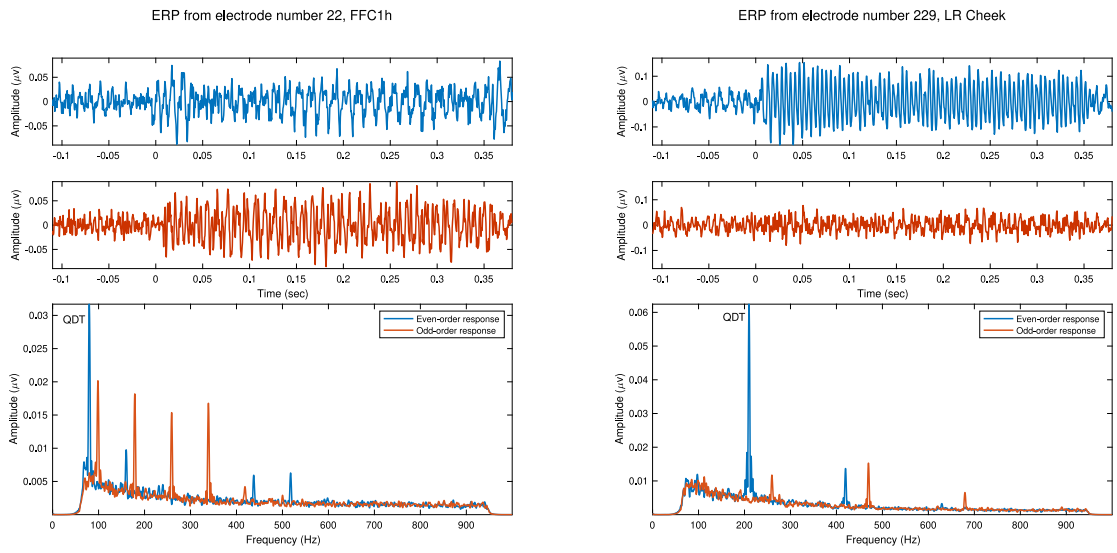


Fig. 2. Example average scalp-space FFRs for low-f0 (A) and high-f0 (B) stimuli. Time-domain waveforms are averaged in the time domain; spectra are averaged spectra. *Y*-axis limits are different to enhance detail, but should be noted. The specific electrode chosen for display for each stimulus was based on which produced the highest SNR of the QDT.

orientation conditions in the spectral domain. All frequency analysis was done as the magnitude of a Fourier transform.

3. Results

3.1. Summary

Source analysis was performed using scalp-space EEG data. But before this was done, the scalp-space EEG data was analyzed to compare it with known aspects of the FFR, and confirm that it was of sufficient integrity for further analysis. Electrodes were averaged over subjects in both the time and spectral domains, in case responses were at different phases for different subjects. Time-domain averages in scalp space were done for visualization (Fig. 2), but spectral averages were analyzed for their frequency content, rather than frequency-transforming the time-domain averages. The even-order portion of the responses showed robust amplitudes in many electrodes at predicted frequencies such as the QDT and its first harmonic. Additionally, in the frequency-domain average, spectral peaks were observed at other second-order nonlinearities, namely summation combination tones of the stimulus primaries (Fig. 2A).

Source analysis was done by creating an inverse model, and then doing a frequency analysis of each of the seven scout time series, for every subject, for every stimulus, for the even-order portions of the responses. Scouts were averaged across subjects in the spectral domain. For the lower-frequency stimuli, there was amplitude at the QDT (80 Hz) from both primary auditory cortices, which was significantly greater than control cortical scout spectra. There was a trend for the left auditory cortical QDT amplitude to be greater than right auditory cortex. Frequency analysis of subcortical scouts showed robust responses at multiple FFR frequencies, such as the QDT and its first harmonic. The amplitudes of the QDT increased as the subcortical auditory system was ascended, an effect which was also significant; thus the MGB had more amplitude than the IC, which had more than the CN.

Source analysis for the higher-f0 stimulus (210 Hz) showed similar but slightly different results than the lower-f0 stimulus. For the higher-f0 stimulus, there was amplitude at the QDT in all four cortical scouts, including the control scouts; however, the amplitude of the QDT in both auditory cortices was significantly greater than in both control cortical regions. There were no other amplitude peaks in the cortical scouts. The subcortical scouts showed robust amplitude at the QDT and its first harmonic.

Upon initial data analysis, a trigger-related artifact was found during the first 60 ms of the even-order portion of most trials, in most subjects. The intensity of the artifact varied across electrodes. Additionally, there was evidence of stimulus-related artifact. These artifacts were not fully resolved during data collection, however two analysis methods were utilized to minimize their effects. The first analysis method to minimize the effect of the trigger artifact was simply not to analyze the first 60 ms. Thus, for source-space data, a Tukey window was utilized that zeroed out the contaminated beginning of each ERP before frequency analysis. In the scalp-space data, a second method of analysis was utilized for data visualization. It was found that a simple PCA captured both artifacts very well in the first two principal components of each ERP matrix. Across all stimuli and subjects, the first two components combined explained an average variance of 72.23% ($SD = 19.22\%$). Thus these first two components were removed from all data. Additionally, there were between three and five samples immediately before and after stimulus onset in the data that were clearly errant and which survived the PCA reduction. For data visualization, those samples were replaced with low-level Gaussian noise in all scalp-space data.

3.2. Scalp-space FFRs

EGI's 256-electrode net is arranged such that sensors are placed along geodesic curves around the head. This system of sensor placement is different than the commonly-used 10/20 system; however, labels of sensor locations based on 10/20 (e.g. Cz, F3, P2) are both familiar and convenient. Such a system of labels and scalp landmarks for a dense EEG net would use 5% increments along the circumference of the head and would thus be called a 10/5 system, as 10/20 and 10/10 use 20% and 10% increments respectively. A 10/5 system has been proposed (Oostenveld and Praamstra, 2001) and refined (Jurcak et al., 2007), however a transformation from a 10/5 system to EGI's Geodesic Sensor Net has not been done. It is desirable for efficient scientific communication to obtain these sensor location approximations, thus this was done manually prior to scalp-space analysis and is depicted in Figs. 1C and 1D.

For scalp-space data visualization and to confirm that a reliable FFR was appearing across subjects, all time-domain electrode data were averaged over subjects, for each stimulus condition. Additionally, spectra were calculated for each electrode separately and averaged across subjects in the spectral domain. A simple measure of signal-to-noise ratio (SNR) was calculated for each electrode's spectrum to choose electrodes for display. The ratio of the QDT peak amplitude and the average noise floor of the summation spectrum (Skoe et al., 2022), excluding frequency ranges outside the main filter passband as well as potential QDT response frequencies and their first harmonics, was calculated as the SNR for each electrode, and the highest was chosen from the large shift condition for plotting. These FFRs are depicted in Fig. 2. A prominent QDT is present in both the low- and high-f0 conditions.

The electrodes selected for display based on QDT SNR exhibit a noticeable difference between the low- and high-f0 stimuli. Electrodes around the mastoid have higher SNRs for the high-f0 stimuli, while the low-f0 stimuli are more variable, but the higher SNRs cluster more on the top, anterior portion of the head. To more completely visualize this pattern, topographical maps of the scalp were constructed to summarize the distribution of the QDT amplitude around the head. Figs. 3A and 3B show these distributions for the low-f0 stimuli, and Figs. 4A and 4B show them for the high-f0 stimuli. The patterns in these scalp maps are consistent with the automated electrode choices by SNR. The locations around the mastoids were important for the high-f0 stimuli. This location is often used for a reference electrode in simple FFR studies because of the mastoid's proximity to the first synapses of the auditory system. In this average-referenced data, it is clear also that this is a site of high-frequency FFR activity, whereas peaks at the low-f0 QDT are also distributed around the top and front of the head, perhaps indicating more spatially-dispersed sources. Additionally, the use of an average reference may tend to enhance signals from electrodes on the periphery of the scalp and suppress central signals; therefore the presence of low-f0 QDTs around the top and front of the head is especially noteworthy.

3.3. Source analysis of FFR

Source analysis of the FFR indicates that there is a cortical contribution at both the low-f0 and high-f0 QDTs, and both the left and right auditory cortex ROIs demonstrated the same peaks. Summaries of all even-order FFRs in source space, arranged with an emphasis on comparing cortical scouts with each other, and subcortical scouts with each other, are contained in Figs. 3C, 3D, 4C and 4D. There was no activity from the control cortical sources (frontal and occipital poles) for the low-f0 QDT in either shift condition, which is clear from Figs. 3C and 4D.

The limits of the y-axes are the same across all even-order plots for easy comparison. Figs. 3E and 3F show a spectrum from the subcortical scouts with the same frequencies as the high-SNR scalp-space FFRs, and around twice the energy at the QDT relative to the auditory cortex sources. For the subcortical scouts, the QDT and its first harmonic (160 Hz) are apparent for the low-f0 stimuli. MGB amplitude is the greatest of the subcortical sources, with IC in the middle and CN being the lowest. This pattern is consistent across all subcortical source analysis.

To fully quantify the differences between all conditions, two main three-way repeated-measures ANOVAs were carried out, because all subjects received all factors and levels and there were no missing data points. First, the dependent variable for these two analyses was the amplitude of the f0 (i.e., QDT frequency) in all spectra (which was the amplitude of 80 Hz component in the case of the low-f0 conditions, and 210 Hz in the case of the high-f0 stimuli). One ANOVA was done for the cortical ROIs, and one for the subcortical ROIs. The three factors were f0 (with two levels: low and high, corresponding to 80 and 210 Hz), shift condition (with two levels: small and large), and ROI (with four levels for cortex: occipital, frontal, left auditory, and right auditory, and three levels for subcortical: MGB, IC, and CN). For both ANOVAs, Mauchly's test of sphericity (Mauchly, 1940) was conducted for the factor ROI because it had more than two levels, and assumptions of equality of variances of differences were not violated (all $p > 0.05$). The ANOVAs are summarized in Figs. 5A and 5B.

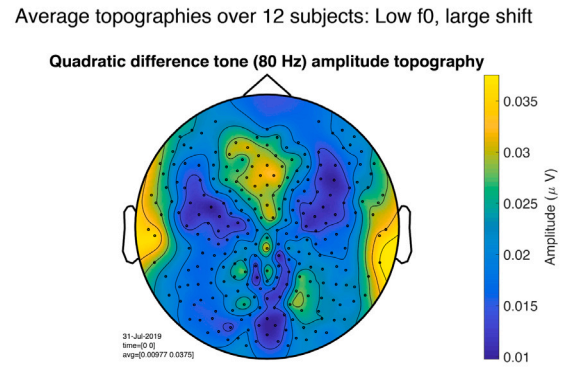
Table 2
Table for cortex ROI model: Three-way repeated measures ANOVA with dependent variable QDT peak amplitude.

Effect	DFn	DFd	F stat	p value	sig.	η_g^2
f0	1	11	6.493	2.70e−02	*	0.090000
shift	1	11	10.280	8.00e−03	*	0.007000
ROI	3	33	18.858	2.65e−07	*	0.262000
f0 x shift	1	11	0.439	5.21e−01	ns	0.000166
f0 x ROI	3	33	5.037	6.00e−03	*	0.077000
shift x ROI	3	33	2.596	6.90e−02	ns	0.007000
f0 x shift x ROI	3	33	0.404	7.51e−01	ns	0.002000

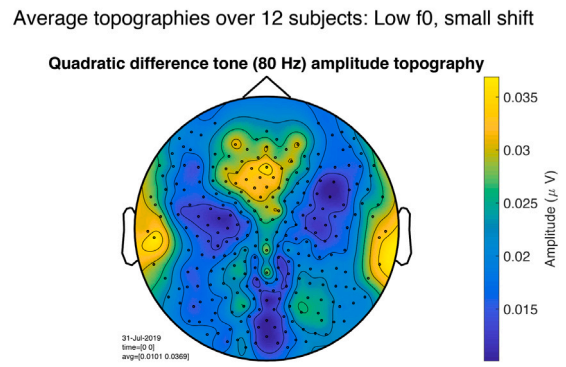
For the cortex ANOVA, all three main effects were significant (Table 2). The main effect of f0 was significant ($F(1,11) = 6.5, p < 0.05$), with low f0 having higher amplitudes than high. The largest effect was the differences between ROIs ($F(3,33) = 18.9, p < 1 \times 10^{-6}$). This is apparent in Figs. 3C, 3D and 5A. As there were significant effects, a full post-hoc pairwise analysis was run with Bonferroni correction. All comparisons between auditory ROIs (left and right auditory cortex) and control ROIs (frontal and occipital) for both f0 conditions show that auditory ROIs have significantly greater response at the QDT than controls (Table 3). There is also a trend for the left auditory cortex response to be greater than the right ($t(23) = 2.6, p = 0.09$) for the low f0 stimulus, a lateralization which is opposite of the significant findings in both Coffey et al. (2016) and Gorina-Careta et al. (2021), though it is not significant here. There was also a significant interaction between f0 and ROI ($F(3,33) = 5.04, p = 6 \times 10^{-3}$), such that overall amplitude was less different between ROIs for the high f0 as compared with the low f0.

The same ANOVA design was run for the subcortical ROIs as well, summarized in Fig. 5B. There was a significant main effect of ROI (Table 4, $F(2,22) = 21.2, p = 7{\times}10^{-6}$), thus a complete post-hoc pairwise analysis was run with Bonferroni correction for multiple comparisons, summarized in Table 5. There were no other significant main effects or interactions. It is apparent in visualizations such as Figs. 3E and 3F that QDT amplitude goes up as the ROIs ascend the subcortical auditory system, and the post-hoc comparisons show that this is a significant effect for all pairwise comparisons between ROIs. The same pattern of relative amplitudes (CN<IC<MGB) in the subcortical scouts is evident for the high-f0 stimuli in Figs. 4E and 4F.

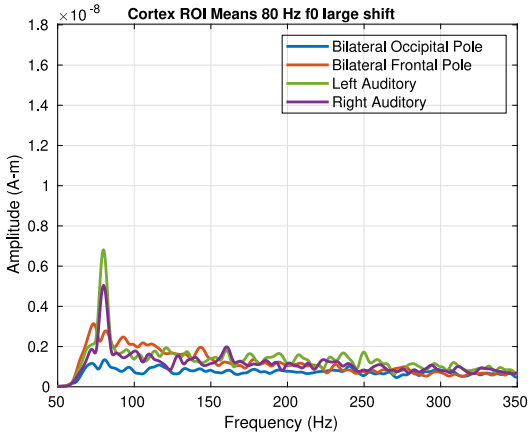
To take into consideration possible effects of differences in noise floor between the cortical ROIS, signal-to-noise ratios (SNRs) were calculated and compared between the auditory and control regions. Spectra for the occipital and frontal control ROIs were averaged, then spectra for the left and right auditory ROIs were averaged, and then SNRs were calculated in the same manner as above for the scalp-space electrode visualizations: QDT frequency amplitudes were divided by the averaged noise floor of the spectrum (Skoe et al., 2022), excluding frequency ranges outside the main filter passband as well as potential QDT response frequencies and their first harmonics. The median SNRs for the low-f0 stimuli were 3.335 and 5.544 for the control and auditory ROIs, respectively. The median SNRs for the high-f0 stimuli were 3.0891 and 3.508 for the control and auditory ROIs.



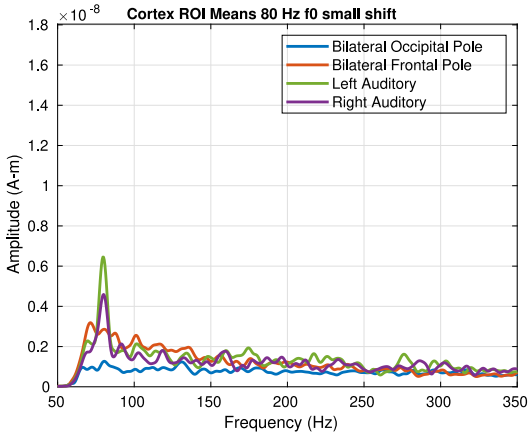
(**A**) Topographies for large shift condition.



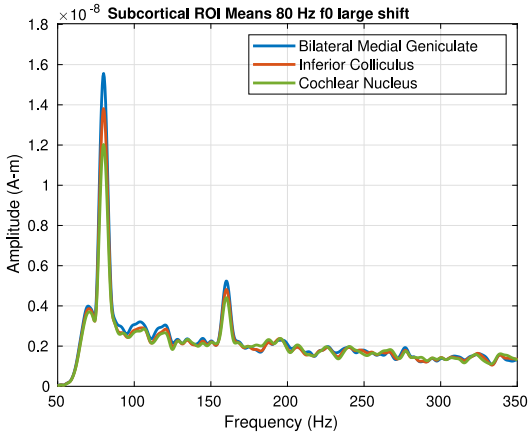
(**B**) Topographies for small shift condition.



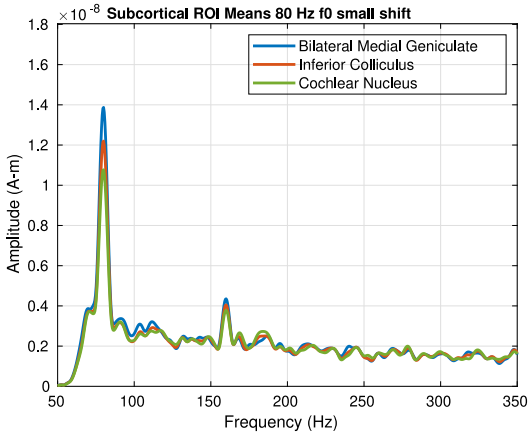
(**C**)



(**D**)



(**E**)



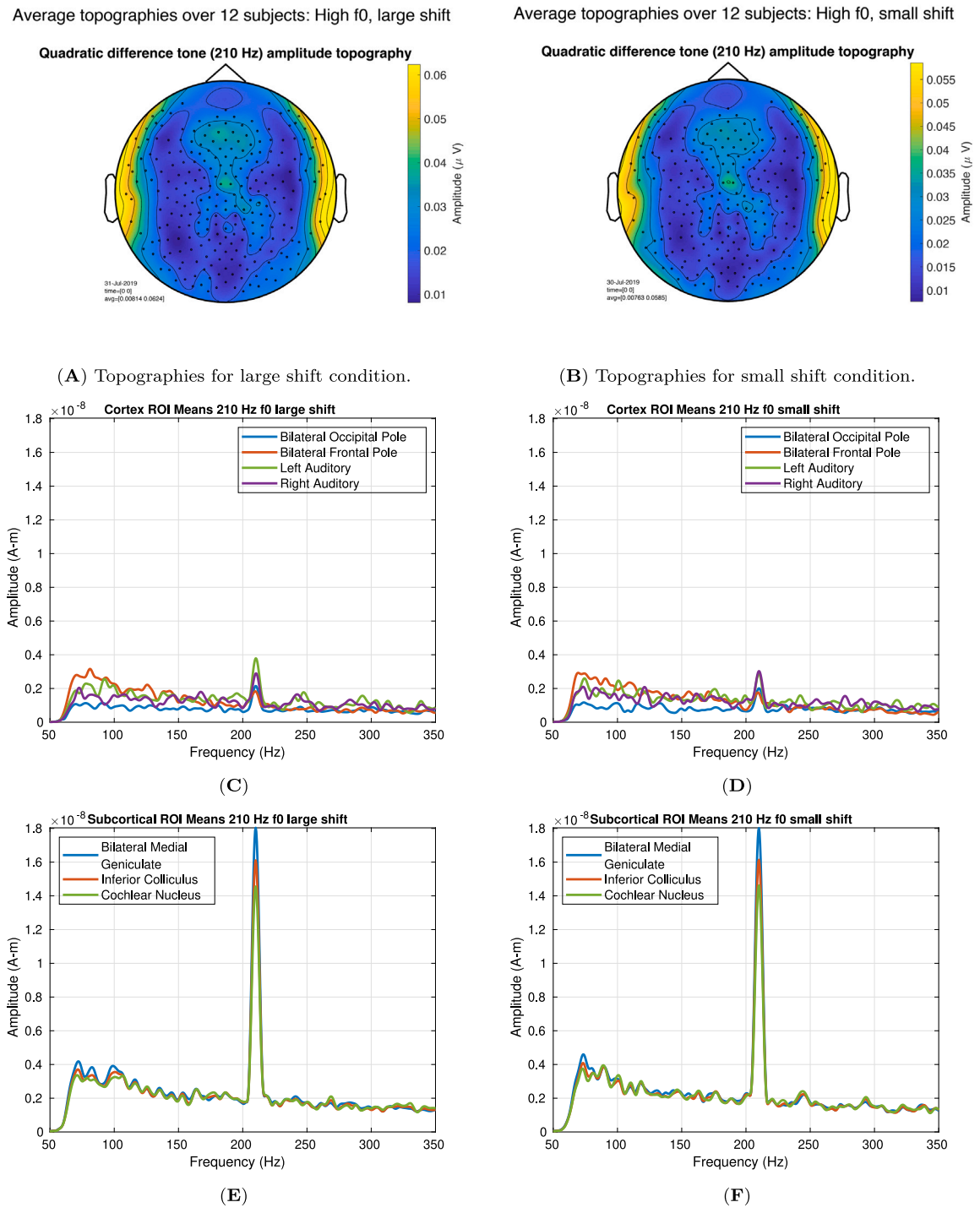
(**F**)

Fig. 3. A–B: Topographies of auditory nonlinearities averaged over subjects for 80 Hz f0 condition. C–D: Cortical source spectra for even-order FFRs in the low-f0 condition. E–F: Subcortical source spectra for even-order FFRs in the low-f0 condition. All source current units are ampere-meters (A-m) by default in Brainstorm's output (Tadel et al., 2011).

Another three-way repeated-measures ANOVA was carried out to analyze these summary SNRs, shown in Fig. 6. As ratios are not in general normally distributed, SNRs were transformed with the natural logarithm before statistical analysis. The main effect of ROI (now Control and Auditory) was again significant (Table 6, $F(1, 11) = 60.193$, $p = 8.73 \times 10^{-6}$) with Auditory greater than Control. Both pairwise posthoc comparisons were also significant (Table 7; Auditory is greater than Control: $t(23) = 7.33$, $p = 1.85 \times 10^{-7}$ for the low-f0 stimuli, and $t(23) = 4.15$, $p = 3.91 \times 10^{-4}$ for the high-f0 stimuli).

4. Discussion

This study presented a detailed source analysis of the frequency following response to both low- (80 Hz) and high-fundamental (210 Hz) auditory stimuli, showing that both types of stimuli elicit FFRs originating from both auditory cortex and brainstem structures. For both low and high f0 stimuli, most of the phase-locked energy that was detected at the scalp originated from subcortical sources, with a small but nonnegligible amount of phase-locked energy from auditory cortical



(**A**) Topographies for large shift condition.

(**B**) Topographies for small shift condition.

(**C**)

(**D**)

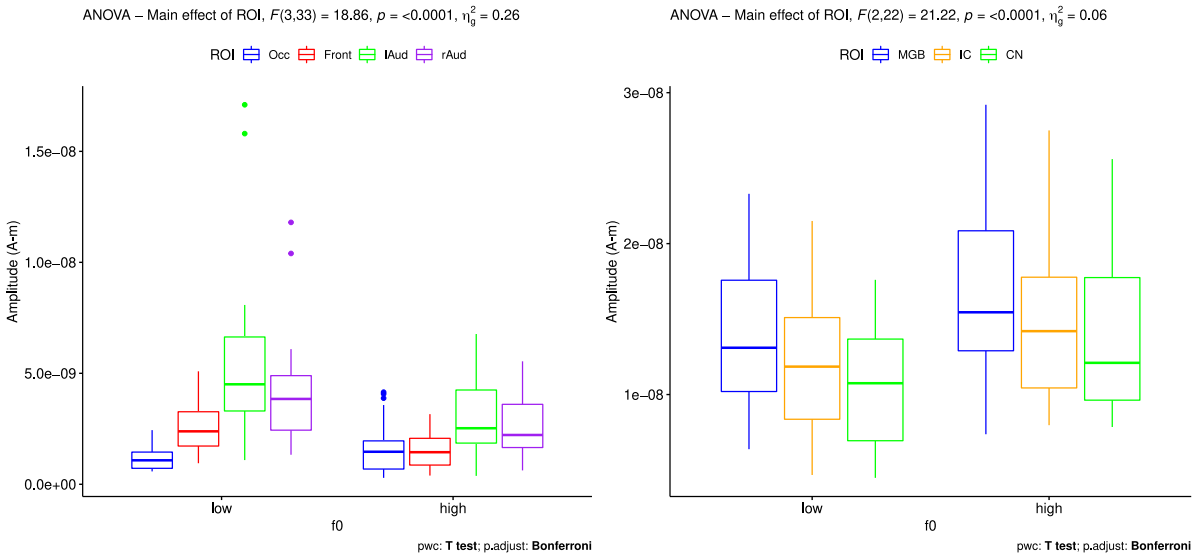
(**E**)

(**F**)

Fig. 4. A–B: Topographies of auditory nonlinearities averaged over subjects for 210 Hz f0 condition. C–D: Cortical source spectra for even-order FFRs in the high-f0 condition. E–F: Subcortical source spectra for even-order FFRs in the high-f0 condition. All source current units are ampere-meters (A-m) by default in Brainstorm's output (Tadel et al., 2011).

regions. While auditory cortex was found to be involved in generating the response to both f0s, there was more cortical source amplitude in response to the low-f0 stimuli than the high-f0 stimuli. Thus, for the low-f0 stimuli, the aggregate response was more heavily weighted by auditory cortical sources than for the high-f0 stimuli. The control cortical regions of interest show significantly less response to the f0 than both auditory cortex ROIs. The latter result held even when accounting for potential differences in noise floors between auditory and control ROIs, as shown through a signal-to-noise ratio (SNR) analysis. The use of two different stimuli for each f0 (i.e., two low f0 and two high f0 stimuli) allowed for findings to be confirmed across multiple stimuli, and the use of missing f0 stimuli ensured that the auditory periphery and brain were responsible for generating the frequencies of interest (QDT). This is the first study to conduct FFR source analysis of this type utilizing EEG – collectively, the constellation of findings suggests that the MEG source imaging results presented by Coffey et al. (2016) and Gorina-Careta et al. (2021) can largely generalize to EEG, the more common approach to measuring the FFR.

The multi-source characteristic of the FFR shown in the source analysis presented here is also evident in single-electrode recordings



(**A**) Cortex ROIs. (**B**) Subcortical ROIs.

Fig. 5. Three-way repeated measures ANOVA results comparing QDT frequency amplitudes. This visualization collapses across the two shift conditions. The four outlying QDT amplitude data points for the auditory cortex ROIs are all from the same subject. For plotting purposes, outliers were defined within the utilized R functions to be higher than 1.5 times the interquartile range above the 75th percentile, and lower than 1.5 times the interquartile range below the 25th percentile. *Y*-axis limits for A are kept lesser than B to show detail, but the difference should be noted.

Table 3
Table for cortex ROI model: All Bonferroni-corrected pairwise comparisons.

f0	ROI1	ROI2	n1	n2	t stat	df	*p* value	*p*.adj	sig.
Low	Occ	Front	24	24	−6.15	23	0.00000284	0.000017	****
Low	Occ	lAud	24	24	−6.06	23	0.0000035	0.000021	****
Low	Occ	rAud	24	24	−6.09	23	0.00000326	0.0000196	****
Low	Front	lAud	24	24	−4.21	23	0.000336	0.002	**
Low	Front	rAud	24	24	−3.62	23	0.001	0.009	**
Low	lAud	rAud	24	24	2.63	23	0.015	0.09	ns
High	Occ	Front	24	24	0.227	23	0.823	1	ns
High	Occ	lAud	24	24	−4.62	23	0.000119	0.000714	***
High	Occ	rAud	24	24	−4.19	23	0.000349	0.002	**
High	Front	lAud	24	24	−4.76	23	0.0000841	0.000505	***
High	Front	rAud	24	24	−3.41	23	0.002	0.014	*
High	lAud	rAud	24	24	1.01	23	0.322	1	ns

Table 4
Table for subcortical ROI model: Three-way repeated measures ANOVA with dependent variable QDT peak amplitude.

Effect	DFn	DFd	F stat	*p* value	sig.	η_g^2
f0	1	11	1.994	1.86e−01	ns	7.30e−02
shift	1	11	2.574	1.37e−01	ns	4.00e−03
ROI	2	22	21.216	7.35e−06	*	6.00e−02
f0 x shift	1	11	4.347	6.10e−02	ns	4.00e−03
f0 x ROI	2	22	0.005	9.95e−01	ns	6.01e−06
shift x ROI	2	22	3.033	6.90e−02	ns	1.48e−04
f0 x shift x ROI	2	22	0.574	5.72e−01	ns	3.41e−05

made to a range of stimulus frequencies. Tichko and Skoe (2017) recorded FFRs to 129 different stimulus f0s between ≈16 and 8000 Hz. Across this stimulus set, the authors found a nonmonotonic and surprisingly richly-structured curve, when the amplitude of the response to the f0 was compared across frequency. They observed that two stimuli close in frequency can yield considerably different response amplitudes, even with the stimulus intensity matched. Frequencies that produced low intensity responses were interpreted as being at least in part due to phase cancellation at the scalp from multiple sources, an idea that they tested both empirically and from a modeling perspective. The two stimulus f0s that were selected for the current work were chosen because they aligned with frequencies that produced very strong responses in the Tichko and Skoe (2017) dataset, and where their model predicted phase alignment across multiple sources.

A comparison between low- and high-f0 stimuli in the current study shows both scalp-space and source-space distinctions in the FFR for these two stimuli. The topography comparisons suggest that most activity for the high-f0 conditions is coming from the brainstem with minimal cortical activity, while the low-f0 responses, in addition to their contributions from subcortical activity, are partially generated by comparatively greater cortical activity relative to the high-f0 condition. And indeed, the source analysis bears this out; Figs. 3C and 3D demonstrate a cortical contribution to the FFR at the component of interest (the QDT) for the low-f0 stimuli. Figs. 4C and 4D show a small but significant contribution from cortex for the high-f0 responses as well (relative to control regions), but with the ratio of cortical-to-subcortical contribution being much smaller for the responses to those stimuli. While additional statistical tests could be carried out comparing cortical and subcortical amplitudes directly, the differences between the modeled source vertices of cortex and those of subcortical volumes would make such a comparison less interpretable; while the cortical sources are constrained for orientation, the volumetric sources are not. In addition to this, the signal-to-noise ratios are likely different between cortical and volumetric sources, an issue that could possibly addressed more in future work in order to more directly compare cortical and subcortical source activity.

Table 5
Table for subcortical ROI model: All Bonferroni-corrected pairwise comparisons.

f0	ROI1	ROI2	n1	n2	t stat	df	*p* value	*p*.adj	sig.
Low	MGB	IC	24	24	7.11	23	0.000000303	0.000000909	****
Low	MGB	CN	24	24	7.15	23	0.000000281	0.000000843	****
Low	IC	CN	24	24	5.09	23	0.0000377	0.000113	***
High	MGB	IC	24	24	4.16	23	0.00038	0.001	**
High	MGB	CN	24	24	4.96	23	0.0000512	0.000154	***
High	IC	CN	24	24	5.14	23	0.0000332	0.0000996	****

Table 6
Table for cortical ROI SNR model: Three-way repeated measures ANOVA with dependent variable $\log_e$ QDT SNR.

Effect	DFn	DFd	F stat	*p* value	sig.	η_g^2
f0	1	11	8.653	1.30e−02	*	0.176000
shift	1	11	9.809	1.00e−02	*	0.019000
ROI	1	11	60.193	8.73e−06	*	0.255000
f0 x shift	1	11	0.189	6.72e−01	ns	0.000260
f0 x ROI	1	11	3.805	7.70e−02	ns	0.026000
shift x ROI	1	11	1.007	3.37e−01	ns	0.004000
f0 x shift x ROI	1	11	0.119	7.37e−01	ns	0.000505

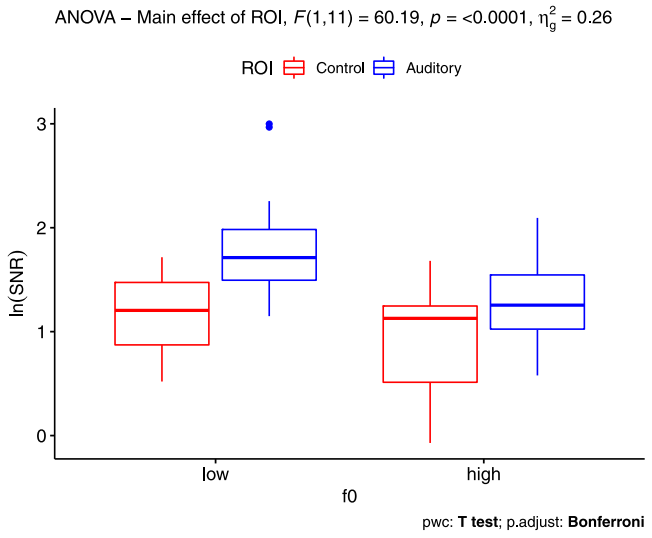


Fig. 6. Three-way repeated measures ANOVA results comparing QDT frequency signal-to-noise ratios (SNRs) in cortical ROIs. Occipital and frontal control ROIs have been averaged in the spectral domain, as have left and right auditory cortical ROIs. SNRs have been transformed with the natural logarithm.

Whereas the amplitude ratio of cortical and subcortical FFR sources for the low-f0 conditions was roughly $\frac{1}{2}$, the ratio for the high-f0 conditions is closer to $\frac{1}{5}$, indicating a much weaker cortical contribution to the FFR, relative to the subcortical contribution, for the higher-frequency stimuli. This is intuitive, but not necessarily obvious. While it is true that most neurons in neocortex do not have fast enough refractory periods to fire an action potential once per cycle at e.g. 210 Hz, it is important to remember that they may still phase lock at that frequency albeit not to every cycle. High phase-locking values require only consistent relative phase between neuronal activity and stimulus, not that there are as many postsynaptic potential peaks as there are stimulus cycles. Thus if many neurons of a population are phase-locking to an input frequency and each is firing at some subset of all input cycles, a population-level response such as the FFR will still show power at the input frequency. This is sometimes called volley theory. In the case of the missing-f0 stimuli used in the current work, the frequency that elicits a phase-locked response, namely the QDT, is a nonlinearity generated at lower levels of the auditory system and not an input frequency that is physically in the stimulus, but the same principles apply to both missing and non-missing-f0 stimuli.

The data from the present study suggest that in addition to the cochlear nucleus and inferior colliculus, the two prominent generators of the FFR are the auditory thalamus (medial geniculate body, MGB) and primary auditory cortex. In addition to and related to this, an interesting difference between the present study and Gorina-Careta et al. (2021) is that the latter only found minimal contribution from the MGB (relative to the CN), whereas the EEG data here suggests that the MGB is the dominant generator of the FFR for both stimulus f0s.

The results of this study also differ from previous studies (Coffey et al., 2016, Gorina-Careta et al., 2021) with respect to cortical lateralization. In general, one three-way repeated-measures ANOVA shows that auditory cortex regions in both hemispheres are responsible for contributions to the FFR (specifically the QDT) to a significantly greater extent than both control cortical regions. In addition to the main effect of ROI, corrected post-hoc tests show that all ROI-to-ROI pairings bear out this relationship (Table 3). There is a trend for a left lateralization among the auditory cortex ROIs, though this is not significant. It should be noted that this is the opposite of the right FFR lateralization, found in both Coffey et al. (2016) and Gorina-Careta et al. (2021), which is sometimes compared to responses corresponding to musicianship and fine pitch discrimination (Coffey et al., 2016). However there were important differences in the types of stimuli used between those two studies and the current study. Coffey et al. (2016) used a speech sound, where Gorina-Careta et al. (2021) used single sinusoids. The current study utilized complex tones comprised of three sinusoids that, when presented together, elicit a response at a lower frequency, the QDT, that is not in the stimulus itself. There is also a main effect of f0, showing that cortical contributions to the FFR are significantly greater for the lower-f0 stimuli. This is consistent with a well-understood dropoff in the ability of cortical neurons to reproduce higher frequencies compared to brainstem neurons.

In another three-way repeated-measures ANOVA, contributions from brainstem sources are examined; another main effect of ROI is observed, with corrected post-hoc tests showing that QDT FFR amplitude significantly increases as the brainstem auditory system is ascended, with thalamus being greater than inferior colliculus, and inferior colliculus being greater than cochlear nucleus.

Utilizing a third repeated-measures ANOVA, cortical auditory ROIs are shown to exhibit greater FFR contributions than control ROIs even when accounting for potential differences in noise floors (Fig. 6, Table 6), which theoretically could artificially increase peak amplitudes. The significantly greater auditory SNRs hold for both the low- and high-f0 stimuli separately (Table 7). While the lack of FFR QDT frequency peaks is apparent in control ROIs for the low-f0 stimuli (Figs. 3C and 3D), there may be small QDT peaks in the control ROIs above the noise floor for the high-f0 stimuli (Figs. 4C and 4D). While both the peak-amplitude and SNR ANOVAs show that cortical auditory ROIs contribute significantly more to the FFR QDT than cortical control ROIs, the presence of the QDT in the control ROI spectra may result either from inverse model source leakage, or a real but weak brain response from those locations, a possible but less likely reason.

Early brainstem sources of the FFR such as the cochlear nuclei generate potentials in lateral directions from a medial part of the head (Starr and Hellerstein, 1971; Starr and Squires, 1980), which may in turn lead to source-model dipoles at parts of cortex oriented orthogonal to those locations and directions (Grandori, 1986), including some

Table 7
Table for cortical ROI SNR model: Both significant pairwise comparisons.

f0	ROI1	ROI2	n1	n2	t stat	df	*p* value	*p*.adj	sig.
Low	Control	Auditory	24	24	−7.33	23	0.000000185	0.000000185	****
High	Control	Auditory	24	24	−4.15	23	0.000391	0.000391	***

that may be errant, such as the more medially-located and vertically-oriented cortical control regions selected in the present study (occipital and frontal poles). As earlier synapses of the auditory system can likely reproduce higher frequencies relative to later synapses, this effect may be more pronounced for responses of higher frequencies, such as the higher-f0 stimuli here. The greater levels of auditory ROI QDT observed here even for the higher-f0 stimuli likely indicate real auditory cortex contribution to the FFR over and above the potential inverse model source errors, but further study of this issue is nevertheless warranted. An extension of the present source analysis (which utilized only dSPM minimum norm imaging (Dale et al., 2000) in Brainstorm) including further modeling utilizing several different inverse model methods and measures would paint a clearer picture of the nature of the activity detected in the control cortical ROIs: If the small FFR frequency peaks observed here are due to errant source modeling, there should be much more variability between source analysis methods for those ROIs than for the auditory cortex ROIs. Various source modeling approaches are readily available, including from within the Brainstorm software; therefore comparing these options is a necessary step for continuing to interpret and understand FFR source analysis in the near future.

While similar studies have been undertaken in recent years, this is the first study to utilize high-density EEG and structural MRI to conduct a source analysis on the FFR to multiple complex tones. The results provide more converging evidence of a cortical contribution to the FFR under most electrode montage regimes, as was also prominently concluded by Coffey et al. (2016) and Gorina-Careta et al. (2021). However, there are multiple aspects of this study that need to be replicated and expanded upon. Firstly, the focus of the current work was on the source of the QDT response, a response that other authors refer to as the envelope following response. The QDT is the dominant auditory response frequency produced by complex harmonic stimuli, but it is not the only nonlinearity in the FFR, and consideration of these other components is necessary for a more comprehensive understanding not just of the FFR's frequency content but also the generators of different components. This can be much more readily achieved with lower impedances than those used out of necessity in the present study. The current work used a 256-electrode EEG net that does not require scalp abrasion, and can be applied in less than 10 min. While most FFR studies maintain electrode impedance less than 5 kΩ, such low impedances cannot be achieved with the nets used here, increasing the likelihood of recording artifacts, including the trigger artifact and stimulus artifact measured in the current work. Although a low-impedance, high-density EEG system would be slow to prepare, a replication attempt of the present source analysis results with such a system would be in order. The advantage of using a high-density EEG system is that there is evidence that the inclusion of inferiorly-located electrodes such as those on the cheeks, below the ears, and on the neck in EGI's 256-electrode net contributes strongly to the ability to localize scalp-space signals to deeper structures such as the brainstem (Song et al., 2015). It is an open question what a source analysis excluding these electrode placements would show with regard to the FFR.

While this study used an ROI-based approach to source analysis, similar to Coffey et al. (2016) and Gorina-Careta et al. (2021), various forms of whole-brain analysis would also be in order, to ensure that all possible sources are accounted for. One such approach is Dynamic Imaging of Coherent Sources (Gross et al., 2001), in which not only is the whole brain analyzed for possible contribution to relevant scalp-space responses, but coherence between brain areas is analyzed to correct for possible volume conduction leakage between sources. Such a whole brain analysis will also shed further light on FFR sources.

The chief limitations of the present study are clear: High impedances, the presence of trigger artifact, and the presence of stimulus artifact. While the impedance problem is difficult to solve if one wants a dense sampling of the scalp space, various artifacts can be vigorously controlled for in future studies. The presence of the trigger artifact unfortunately meant that a crucial part of the analysis in Coffey et al. (2016) could not be replicated, namely the delay-based identification of successive auditory structures. The onset responses from most of the auditory system are over well before 60 ms, which was the length of the trigger artifact in each trial here. Fortunately frequency analysis could still be done on the remainder of the FFR, but more detailed study of the combined onset responses from successive structures is highly desirable for the future.

Despite these limitations, this is the first study to show a detailed MRI-constrained FFR source analysis with EEG, which is a technique that is more robust to detecting deep sources relative to MEG. Source configurations of the FFR are becoming an increasingly common topic of study (Zhang and Gong, 2019; Bidelman and Momtaz, 2021; Gnanateja et al., 2021), thus the present study is timely. While it does replicate certain aspects of previous similar work with MEG, there are also important differences in the current results that need to be explored further. As more is discovered about the frequency following response, it can more readily be utilized in clinical and diagnostic settings. Having a clearer understanding of the neural sources of the FFR could allow for more refined auditory diagnostics, treatments, and monitoring. While there is still much to be learned about the FFR, one characteristic is becoming increasingly clear: The FFR measured from the scalp has multiple sources.

CRediT authorship contribution statement

Karl D. Lerud: Conceptualization, Methodology, Implementation, Analysis, Writing. **Roeland Hancock:** Conceptualization, Methodology, Implementation, Supervision, Writing. **Erika Skoe:** Conceptualization, Methodology, Funding, Supervision, Writing.

Declaration of competing interest

None

Data availability

Code and data are publicly available at: https://osf.io/2ar9q/.

Acknowledgments

This work was supported by a University of Connecticut Brain Imaging Research Center, United States (UConn BIRC) seed grant to KL and National Science Foundation Funding, United States to ES (NSF1941147).

References

Ahlfors, S.P., Han, J., Belliveau, J.W., Hämäläinen, M.S., 2010. Sensitivity of MEG and EEG to source orientation. Brain Topogr. 23, 227–232.

Bhagat, S.P., Champlin, C.A., 2004. Evaluation of distortion products produced by the human auditory system. Hear. Res. 193, 51–67.

Bian, L., Chen, S., 2008. Comparing the optimal signal conditions for recording cubic and quadratic distortion product otoacoustic emissions. J. Acoust. Soc. Am. 124, 3739–3750.

Bidelman, G.M., 2015. Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR. Hear. Res. 323, 68–80.
Bidelman, G.M., 2018. Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. NeuroImage 175, 56–69.
Bidelman, G.M., Momtaz, S., 2021. Subcortical rather than cortical sources of the frequency-following response (FFR) relate to speech-in-noise perception in normal-hearing listeners. Neurosci. Lett. 746, 135664.
Chimento, T.C., Schreiner, C.E., 1990. Selectively eliminating cochlear microphonic contamination from the frequency-following response. Electroencephalogr. Clin. Neurophysiol. 75, 88–96.
Coffey, E.B.J., Herholz, S.C., Chepesiuk, A.M.P., Baillet, S., Zatorre, R.J., 2016. Cortical contributions to the auditory frequency-following response revealed by MEG. Nat. Commun. 7, 1–11.
Coffey, E.B.J., Musacchia, G., Zatorre, R.J., 2017. Cortical correlates of the auditory frequency-following and onset responses: EEG and fMRI evidence. J. Neurosci. 37, 830–838.
Coffey, E.B., Nicol, T., White-Schwoch, T., Chandrasekaran, B., Krizman, J., Skoe, E., Zatorre, R.J., Kraus, N., 2019. Evolving perspectives on the sources of the frequency-following response. Nat. Commun. 10 (5036).
Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., Louis, S., 2000. Neurotechnique mapping : Combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26, 55–67.
Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968–980.
Farahani, E.D., Goossens, T., Wouters, J., van Wieringen, A., 2017. Spatiotemporal reconstruction of auditory steady-state responses to acoustic amplitude modulations: Potential sources beyond the auditory pathway. NeuroImage 148, 240–253.
Fischl, B., 2012. FreeSurfer. NeuroImage 62, 774–781.
Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341–355.
Gardi, J., Merzenich, M., McKean, C., 1979. Origins of the scalp recorded frequency-following response in the cat. Audiol. Off. Organ Int. Soc. Audiol. 18, 358–381.
Gnanateja, G.N., Maruthy, S., 2019. Dichotic phase effects on frequency following responses reveal phase variant and invariant harmonic distortion products. Hear. Res. 380, 84–99.
Gnanateja, G.N., Rupp, K., Llanos, F., Remick, M., Pernia, M., Sadagopan, S., Teichert, T., Abel, T.J., Chandrasekaran, B., 2021. Frequency-following responses to speech sounds are highly conserved across species and contain cortical contributions. eNeuro 8, 1–20.
Gockel, H.E., Carlyon, R.P., Mehta, A., Plack, C.J., 2011. The frequency following response (FFR) may reflect pitch-bearing information but is not a direct representation of pitch. J. Assoc. Res. Otolaryngol. 12, 767–782.
Gockel, H.E., Farooq, R., Muhammed, L., Plack, C.J., Carlyon, R.P., 2012. Differences between psychoacoustic and frequency following response measures of distortion tone level and masking. J. Acoust. Soc. Am. 132, 2524–2535.
Goldenholz, D.M., Ahlfors, S.P., Hämäläinen, M.S., Sharon, D., Ishitobi, M., Vaina, L.M., Stufflebeam, S.M., 2008. Mapping the signal-to-noise-ratios of cortical sources in magnetoencephalography and electroencephalography. Hum. Brain Mapp. 30, 1077–1086.
Gorina-Careta, N., Kurkela, J.L., Hämäläinen, J., Astikainen, P., Escera, C., 2021. Neural generators of the frequency-following response elicited to stimuli of low and high frequency: A magnetoencephalographic (MEG) study. NeuroImage 231.
Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., Parkkonen, L., Hämäläinen, M.S., 2014. MNE software for processing MEG and EEG data. NeuroImage 86, 446–460.
Gramfort, A., Papadopoulo, T., Olivi, E., Clerc, M., 2010. OpenMEEG: Opensource software for quasistatic bioelectromagnetics. BioMed. Eng. OnLine 9, 1–20.
Gramfort, A., Papadopoulo, T., Olivi, E., Clerc, M., 2011. Forward field computation with OpenMEEG. Comput. Intell. Neurosci. 2011, 1–13.
Grandori, F., 1986. Field analysis of auditory evoked brainstem potentials. Hear. Res. 21, 51–58.
Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., Salmelin, R., 2001. Dynamic imaging of coherent sources: Studying neural interactions in the human brain. Proc. Natl. Acad. Sci. 98, 694–699.
Jurcak, V., Tsuzuki, D., Dan, I., 2007. 10/20 10/10 and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. NeuroImage 34, 1600–1611.
Lerud, K.D., Almonte, F.V., Kim, J.C., Large, E.W., 2014. Mode-locking neurodynamics predict human auditory brainstem responses to musical intervals. Hear. Res. 308, 41–49.
Mauchly, J.W., 1940. Significance test for sphericity of a normal n-variate distribution. Ann. Math. Stat. 11, 204–209.
Moore, G.A., Moore, B.C.J., 2003. Perception of the low pitch of frequency-shifted complexes. J. Acoust. Soc. Am. 113, 977–985.
Nolan, H., Whelan, R., Reilly, R.B., 2010. FASTER: Fully automated statistical thresholding for EEG artifact rejection. J. Neurosci. Methods 192, 152–162.
Nuttall, A.L., Dolan, D.F., 1993. Intermodulation distortion (F2-F1) in inner hair cell and basilar membrane responses. J. Acoust. Soc. Am. 93, 2061–2068.
Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011.
Oostenveld, R., Praamstra, P., 2001. The five percent electrode system for high-resolution EEG and ERP measurements. Clin. Neurophysiol. 112, 713–719.
Schouten, J.F., Ritsma, R.J., Cardozo, B.L., 1962. Pitch of the residue. J. Acoust. Soc. Am. 294, 1418–1424.
Skoe, E., García-Sierra, A., Ramírez-Esparza, N., Jiang, S., 2022. Automatic sound encoding is sensitive to language familiarity: Evidence from English monolinguals and Spanish-English bilinguals. Neurosci. Lett. 777, 136582.
Skoe, E., Kraus, N., 2010. Auditory brain stem response to complex sounds : A tutorial. Ear Hear 31, 1–23.
Smith, J.C., Marsh, J.T., Brown, W.S., 1975. Far-field recorded frequency-following responses - evidence for the locus of brainstem sources. Electroencephalogr. Clin. Neurophysiol. 39, 465–472.
Sohmer, H., Pratt, H., Kinarti, R., 1977. Sources of frequency following responses (FFR) in man. Electroencephalogr. Clin. Neurophysiol. 42, 656–664.
Song, J., Davey, C., Poulsen, C., Luu, P., Turovets, S., Anderson, E., Li, K., Tucker, D., 2015. EEG source localization: Sensor density and head surface coverage. J. Neurosci. Methods 256, 9–21.
Starr, A., Hellerstein, D., 1971. Distribution of frequency following responses in cat cochlear nucleus to sinusoidal acoustic signals. Brain Res. 33, 367–377.
Starr, A., Squires, K., 1980. Distribution of auditory brainstem potentials over the scalp and nasopharynx in humans. Ann. New York Acad. Sci. 338, 427–442.
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.
Teichert, T., Gnanateja, G.N., Sadagopan, S., Chandrasekaran, B., 2022. A linear superposition model of envelope and frequency following responses may help identify generators based on latency. Neurobiol. Lang. 3, 441–468.
Tichko, P., Skoe, E., 2017. Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators. Hear. Res. 348, 1–15.
Zhang, X., Gong, Q., 2019. Frequency-following responses to complex tones at different frequencies reflect different source configurations. Front. Neurosci. 13, 1–18.
Zwicker, E., 1979. Different behaviour of quadratic and cubic difference tones. Hear. Res. 1, 283–292.