OXFORD

Causal Relationship between the Right Auditory Cortex and Speech-Evoked Envelope-Following Response: Evidence from Combined Transcranial Stimulation and Electroencephalography

Guangting Mai[1,2,3] and Peter Howell[3]

[1]Hearing Theme, National Institute for Health Research Nottingham Biomedical Research Centre, Nottingham NG1 5DU, UK, [2]Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham NG7 2UH, UK and [3]Department of Experimental Psychology, University College London, London WC1H 0AP, UK

Address correspondence to Guangting Mai, NIHR Nottingham Biomedical Research Centre, Ropewalk House, 113 The Ropewalk, Nottingham NG1 5DU, UK. Email: guangting.mai@nottingham.ac.uk

Abstract

Speech-evoked envelope-following response (EFR) reflects brain encoding of speech periodicity that serves as a biomarker for pitch and speech perception and various auditory and language disorders. Although EFR is thought to originate from the subcortex, recent research illustrated a right-hemispheric cortical contribution to EFR. However, it is unclear whether this contribution is causal. This study aimed to establish this causality by combining transcranial direct current stimulation (tDCS) and measurement of EFR (pre- and post-tDCS) via scalp-recorded electroencephalography. We applied tDCS over the left and right auditory cortices in right-handed normal-hearing participants and examined whether altering cortical excitability via tDCS causes changes in EFR during monaural listening to speech syllables. We showed significant changes in EFR magnitude when tDCS was applied over the right auditory cortex compared with sham stimulation for the listening ear contralateral to the stimulation site. No such effect was found when tDCS was applied over the left auditory cortex. Crucially, we further observed a hemispheric laterality where aftereffect was significantly greater for tDCS applied over the right than the left auditory cortex in the contralateral ear condition. Our finding thus provides the first evidence that validates the causal relationship between the right auditory cortex and EFR.

Key words: transcranial direct current stimulation (tDCS), envelope-following response (EFR), electroencephalography (EEG), auditory cortex

Introduction

The speech-evoked frequency-following responses (FFRs) are phase-locked neural activities that reflect early processing of periodic features of input speech signals in the human brain (Aiken and Picton 2008; Coffey et al. 2019). One of the most important FFR components in the central auditory systems is the envelope-following response (EFR) that encodes the period-icity envelopes at fundamental frequency (F_0) that represents the vocal pitch information (Aiken and Picton 2008; Coffey et al. 2019). The EFR is associated with various human auditory and language processing. For example, it reflects the neural encoding of linguistic pitch and is stronger in tonal language than nontonal language speakers (Krishnan et al. 2004, 2005, 2009). It has greater strength in musicians who have better pitch

discrimination ability than people without musical training (Musacchia et al. 2007; Wong et al. 2007; Strait et al. 2009; Bidelman et al. 2011). Also, EFR is associated with speech-in-noise perception. Greater EFR magnitudes are associated with better speech recognition ability in noisy environments (Parbery-Clark et al. 2011; Song et al. 2011). EFR also reflects neural plasticity related to fundamental cognitive and physiological processes such as auditory learning (Skoe et al. 2014), changes in arousal (Mai et al. 2019), and attention (Lehmann and Schönwiesner 2014; Hartmann and Weisz 2019).

Clinically, EFR is proposed as a potential biomarker for various auditory and language disorders. EFR declines with age (Anderson et al. 2012; Presacco et al. 2016) and can predict word recognition ability during speech-in-noise perception in older adults (Anderson et al. 2011; Fujihira and Shiraishi 2015; Mai et al. 2018). This indicates that degradations to EFR could potentially explain the increased speech-in-noise difficulty experienced during aging. EFRs are also associated with hearing deficits such as cochlear synaptopathy (Encina-Llamas et al. 2019) and auditory processing disorders (Schochat et al. 2017). Furthermore, deficits of EFR often occurred along with functional impairments in learning and cognitive disorders, such as learning difficulties in literacy (Cunningham et al. 2001; Banai et al. 2007; White-Schwoch et al. 2015), dyslexia (Hornickel and Kraus 2013), and autism (Russo et al. 2008) in children and mild cognitive impairment in older adults (Bidelman et al. 2017).

Because of the relationship between EFR and these fundamental and clinical auditory and language processes, it is essential to understand how different parts of the auditory systems may contribute to EFR. It has long been argued that EFR reflects the encoding of periodicity in the inferior colliculus (IC) at the brainstem, which has been proposed as the main neural source of EFR (Chandrasekaran and Kraus 2010; Bidelman 2015, 2018). Recent studies, however, have shown cortical contributions to EFR (Coffey et al. 2016, 2017a; Hartmann and Weisz 2019; Ross et al. 2020; Gorina-Careta et al. 2021). These studies localized the EFR sources along the auditory pathway using magnetoencephalography (MEG) and illustrated significant neural contributions to EFR within the right auditory cortex associated with musical experience, pitch discrimination ability (Coffey et al. 2016), speech-in-noise perception (Coffey et al. 2017a), intermodal attention (Hartmann and Weisz 2019), and aging (Ross et al. 2020). Another study found that EFR strength was associated with right, but not left, hemispheric hemodynamic activity in the auditory cortex (Coffey et al. 2017b), consistent with the relative specialization of right auditory cortex for pitch and tonal processing (Zatorre and Belin 2001; Patterson et al. 2002; Hyde et al. 2008; Albouy et al. 2013; Cha et al. 2016).

Despite findings that show the potential cortical contribution to EFRs, it is unclear whether the relationship between auditory cortex and EFR is causal and is not simply an observation based on specific source localization techniques or apparent association between cortical activations and EFR. More specifically, it is unclear whether EFR is susceptible to neuroplasticity at the cortical level, for example, how plastic changes in neural excitability of the auditory cortex may cause changes in EFR. An important next step, therefore, is to determine such causality. One way to achieve this aim is to use neurostimulation tools that are known to change cortical excitability and test how EFR may be altered according to these changes. A recent study investigated how EFR obtained via scalp-recorded electroencephalography (EEG) was modulated by inhibiting excitability of the right auditory cortex using repetitive transcranial magnetic stimulation (rTMS)

(López-Caballero et al. 2020). This study, however, did not find significant changes in EFR following stimulation compared with sham. The authors argued that the absence of the aftereffect may be attributed to the ineffectiveness of rTMS to change neural excitability in the auditory sensory region (López-Caballero et al. 2020). Here, we address several critical caveats. First, it is possible that the right auditory cortex contributes to EFR mainly along the contralateral pathway. Stimulation over the right auditory cortex might result in significant changes in EFR when participants listen to sounds from the contralateral ear (i.e., the left ear from which the right auditory cortex receives the majority of its auditory information), rather than the ipsilateral ear (i.e., the right ear). If this is the case, binaural listening during EFR measurements in López-Caballero et al. (2020) might smear the rTMS effect dominant in the contralateral ear. Therefore, using paradigms of monaural listening that allow studying ipsilateral/contralateral effects of stimulation should provide better insights into cortical contributions to EFR. Second, although previous research suggested a right-hemispheric cortical contribution to EFR (Coffey et al. 2016, 2017a, 2017b; Hartmann and Weisz 2019; Ross et al. 2020), a more rigorous design is to include the left auditory cortex as a control stimulation site. This could determine whether cortical contribution is made specifically in the right hemisphere and whether laterality effects occur by comparing aftereffects of stimulations over different hemispheres. Finally, EEG has a poor source localization capability to determine the site of occurrence of the aftereffects. It is therefore necessary to look into the aftereffects at different neural latencies. EFR originates at the auditory brainstem at 5–15 ms after stimulus onset (Chandrasekaran and Kraus 2010). Recent research has further illustrated that there are cortical EFR activities that peak at 50–60 ms after sound onset (Coffey et al. 2016). Looking into timing characteristics of aftereffects could illustrate whether any effects start at the early (brainstem) or later (auditory cortex) auditory centers. Note that "cortical contribution" here we refer to includes not only contributions of cortical EFR, but also corticofugal modulation on the subcortical EFR (Price and Bidelman 2021) even without cortical sources. Examining the timing characteristics could help disentangle whether the cortical contribution is likely made directly at the cortex or through top-down corticofugal modulation on the subcortical level.

Based on the discussions above, the aim of the current study was to establish a causal relationship between the auditory cortex and EFR. Here, we applied transcranial direct current stimulation (tDCS) to alter neural excitability in the left and right auditory cortices. We measured the EFR using scalp-recorded EEG during monaural listening (left and right ears respectively to allow for testing contralaterality) to a repeatedly presented speech syllable pre- and post-tDCS. We then examined the tDCS aftereffects on EFR. tDCS is a noninvasive neurostimulation that modulates cortical excitability (Jacobson et al. 2012). By applying direct currents over the scalp, tDCS leads to neural excitation or inhibition in proximal parts of the cortex that last for up to 90 min poststimulation (Nitsche and Paulus 2001). Previous research has shown that tDCS over auditory cortices can modulate cortical auditory-evoked responses to pitch (recorded via EEG) (Zaehle et al. 2011; Impey and Knott 2015; Boroda et al. 2020). As for hemispheric laterality, previous studies showed that applying tDCS over the right, compared with the left, auditory cortex can significantly change pitch discrimination performances, supporting the causal role of the right auditory cortex for pitch perception (Mathys et al. 2010; Matsushita et al.

2015, 2021). However, such causality has not been established for neurophysiological signatures like EFR. Should cortical contributions to EFR exist, altering cortical excitability via tDCS should cause consequent changes in EFR. The current study tested the hypothesis that tDCS over the right auditory cortex results in changes in EFR that should occur particularly along the contralateral auditory pathway where participants listen to speech from the left ear. Besides, we further examined the hemispheric laterality by directly comparing stimulations between the 2 hemispheres and investigated the timing characteristics of the aftereffects.

Materials and Methods

The present study was approved by the UCL Research Ethics Committee, and informed consents were obtained from all participants who were recruited for the experiments. The recorded EEG data can be publicly accessed through our online repository (https://gin.g-node.org/guangting-mai/tdcs-eeg-data_cercor-bhab298).

Participants

Ninety right-handed, normal-hearing participants (18–40 years old; 45 females) were recruited and completed the entire experiment. Two other participants dropped out during the tDCS phase because they felt uncomfortable with the skin sensation when stimulation was applied. All participants are right-handed (handedness index (HI) > 40; Oldfield 1971) and had normal hearing (pure-tone audiometric (PTA) thresholds ≤ 25 dB HL within the range of 0.25–6 kHz for both ears). Participants were all nontonal language (English, Spanish, Portuguese, Polish, or Russian) speakers, had no long-term musical training, and reported no history of neurological or speech/language disorders. No participants participated in any brain stimulation experiments within the 2 weeks prior to the present experiment.

Participants were assigned at random to 1 of the 5 groups, each of which received different types of tDCS. HI, age, gender, and audiometric thresholds were all matched across the 5 groups (see "tDCS in Experimental Design" for details).

Experimental Design

The experimental procedure is summarized in Figure 1. EFRs were recorded pre- and post-tDCS during monaural listening to a repeatedly presented speech syllable to test for any aftereffects of tDCS. Details of the EFR recording and tDCS protocols are described in the following sections.

Speech Stimulus for the EFR Recording

A 120-ms-long syllable /i/ spoken by a male with a static fundamental frequency (F_0) at 136 Hz was used for the EFR recordings. The reason for using /i/ was that EFR elicited by this stimulus had been proved to be robustly measured in our lab (Mai et al. 2018, 2019). The syllable has a stable amplitude profile across the syllable period except for the 5-ms rising and falling cosine ramps applied at the onset and offset to avoid transients. The waveform of the syllable is shown in Figure 2A.

The syllable stimulus was repeatedly presented at ∼4 times per second with interstimulus interval (ISI, time interval between offset of 1 stimulus and onset of the following stimulus) fixed at 120 ms. The stimulus was played monaurally via electromagnetically shielded inserted earphone (ER-3 insert earphone, Intelligent Hearing Systems, Miami, FL) at 85 dB SPL

(excluding ISIs) in each ear (e.g., left ear listening followed by right ear listening or vice versa with the order of ear presentation counterbalanced across participants). For each ear, there were 3000 sweeps in total with half of the sweeps (i.e., 1500 sweeps) having positive polarity and the other half having negative polarity. The reason for having two different polarities was to minimize neural responses to the temporal fine structure (TFS), which mainly occur at auditory periphery (Aiken and Picton 2008) and electrical artifacts by adding responses to the positive and negative sweeps (see "EEG Signal Processing"). Sweeps with different polarities were presented in an intermixed order for all participants (set prior to the experiments via pseudorandomization).

EFR Recording

EEGs were recorded over participants' scalps using an ActiveTwo system (Biosemi ActiView, the Netherlands) with sampling rate of 16 384 Hz while they listened to the 3000 sweeps of the syllable stimulus both pre- and post-tDCS. Because of the time constraints for which EFR recording needed to be completed as soon as possible after finishing the tDCS phase (in order to maintain the offline effect of tDCS), only 3 active electrodes (Cz, C3 and C4) were placed to save time for EEG setups. The electrodes were localized using a standard Biosemi cap. Here, we focused on Cz only, as it is the standard conventional site used for obtaining robust EFR (Skoe and Kraus 2010). Bilateral earlobes served as the reference and ground electrodes were CMS and DRL at the parieto-occipital sites. Electrode offsets were always kept within ±35 mV.

Participants were seated comfortably in an armchair in an electromagnetically and sound-shielded booth. They listened passively to the stimuli while keeping their eyes on a fixation cross in the center of a computer screen. The 3000 syllable sweeps in each ear were broken into six 2-minute-long blocks (500 sweeps each) with ∼40 s breaks between blocks to minimize fatigue. Participants were required to keep awake and refrain from body and head movements while they were listening to the sounds. Participants were instructed to keep awake because our previous study has shown that EFR magnitudes decrease significantly when participants fall sleep (Mai et al. 2019). Allowing participants to sleep could result in varying levels of arousal that affected the EFR in an uncontrolled manner and therefore could cause confounds for quantifying the tDCS aftereffects. The EFR recording lasted for ∼30 min for both pre- and post-tDCS. The post-tDCS recording was completed within 45 min after tDCS was stopped for all participants to ensure that any aftereffects of tDCS on EFRs were sustained (Nitsche and Paulus 2001).

PTA

Along with EFR recording, audiometric thresholds (PTA at 0.25, 0.5, 1, 2, 3, 4 and 6 kHz) in both ears were also tested pre- and post-tDCS using a MAICO MA41 Audiometer (MAICO Diagnostics, Germany). These were respectively conducted prior to the pre- and post-tDCS EFR recording to test whether tDCS changed peripheral hearing and whether aftereffects on EFR were related to these changes because EFR could be influenced by individual's audibility (Ananthakrishnan et al. 2016). Seven participants' post-tDCS PTA data were not recorded (1 in the Left-AC Anodal group, 2 in the Right-AC Anodal group, 3 in the Right-AC Cathodal group, and 1 in the Sham group; see group allocations in "tDCS") due to the unavailability of equipment at the time of testing. Other than these absent data, PTAs at all frequencies were within the normal range (≤ 25 dB HL) both pre- and post-tDCS in both ears. Subsequent analyses showed
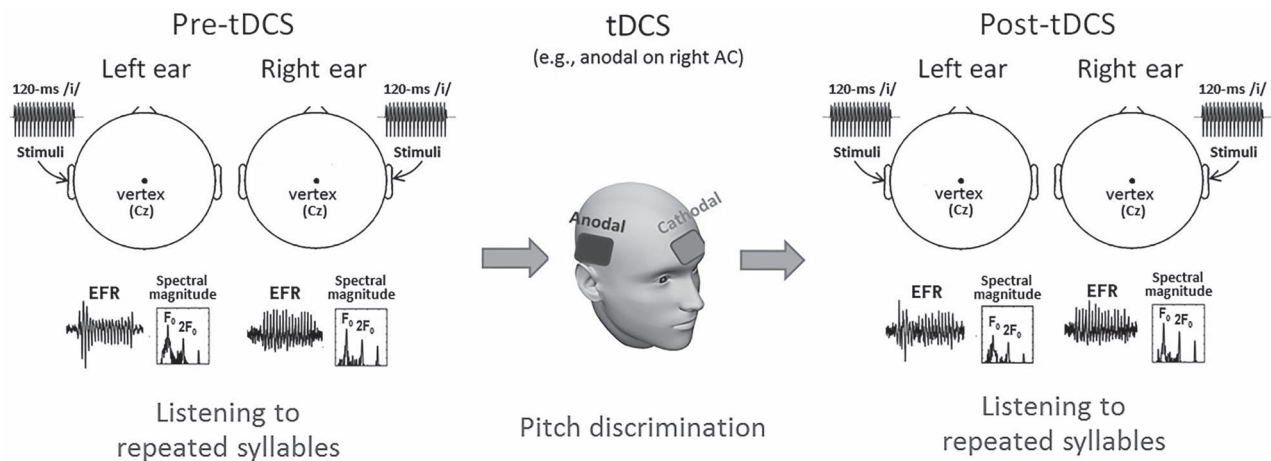
Figure 1. The experimental design. Participants underwent 3 phases (pre-tDCS, tDCS, and post-tDCS). EFRs were measured via scalp EEG at Cz when participants listened to a repeatedly presented syllable /i/ (120 ms long) monaurally pre- and post-tDCS. A pitch discrimination task was performed during the tDCS application over the auditory cortex with the reference electrode placed above the contralateral eyebrow on the forehead.
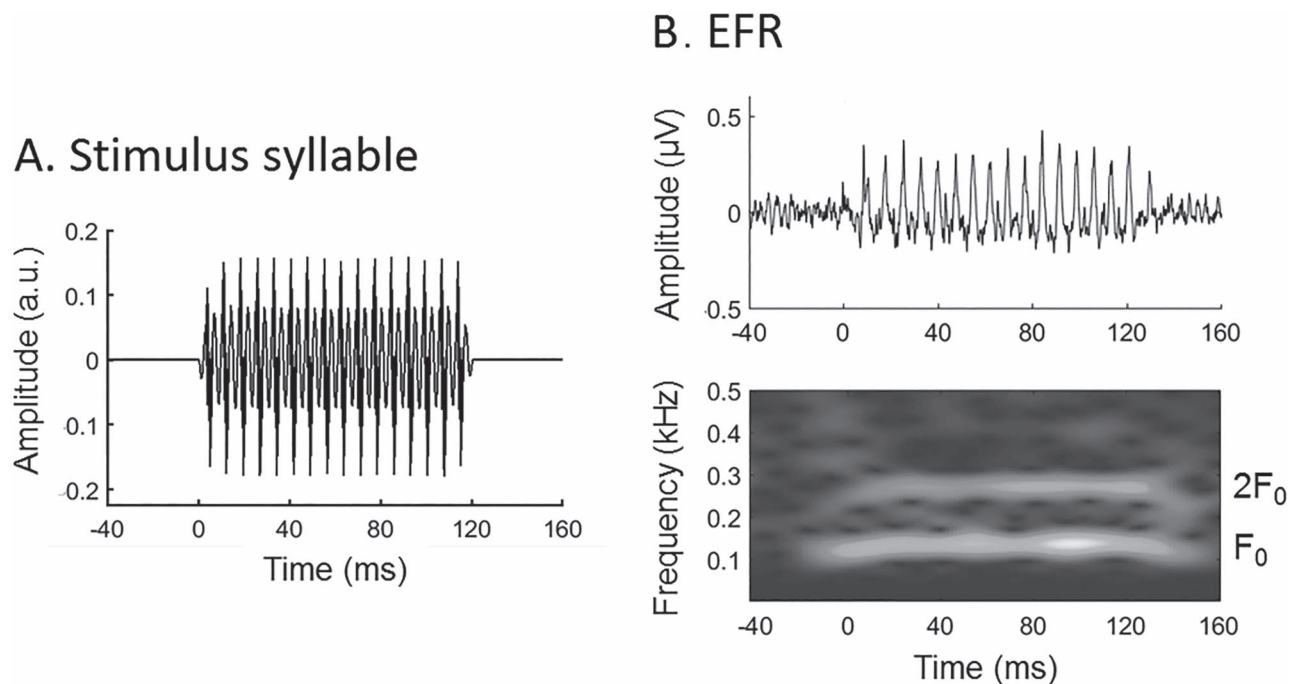

Figure 2. The syllable stimulus for EFR recordings. (A) Temporal waveform of the 120-ms stimulus syllable /i/ with fundamental frequency (F_0) at 136 Hz. (B) A representative sample waveform of EFR (upper) and its spectrogram (lower) showing that the powers are predominated by F_0 and $2F_0$.

baseline (pre-tDCS) PTA and change in PTA (post- vs. pre-tDCS) did not differ significantly between stimulation groups for either ear (see "tDCS" for details). Also no significant association was found between the aftereffect on EFR and baseline PTA or change in PTA (see "Results").

tDCS
tDCS was applied over the scalp using a battery-driven direct current stimulator (Magstim HDCStim, United Kingdom) with a pair of rubber-surface electrodes (5 × 5 cm) contained in saline-soaked cotton pads. Center position of the active electrode was on T7/T8 (according to the 10/20 EEG system) for the left/right auditory cortex. The exact placement of T7/T8 was determined

with the help of a standard Biosemi EEG cap. The reference electrode was placed on the forehead above the eyebrow contralateral to the active electrode (see Matsushita et al. 2015; also see Fig. 1).

Participants were assigned at random to 1 of the 5 groups (18 participants per group; single-blinded). The 5 groups received the following different types of tDCS: 1) anodal stimulation on the left auditory cortex (Left-AC Anodal), 2) cathodal stimulation on the left auditory cortex (Left-AC Cathodal), 3) anodal stimulation on the right auditory cortex (Right-AC Anodal), 4) cathodal stimulation on the right auditory cortex (Right-AC Cathodal), and 5) Sham. There were equal numbers of males and females in each group. Subsequent checks via independent-sample *t*-tests

between groups confirmed that participants' age, PTA (averaged across 0.25–6 kHz for each ear), changes in PTA (post- vs. pre-tDCS for each ear), and HI were matched across groups (all $P > 0.3$, False discovery rate (FDR) corrected according to multiple (10) comparisons between the 5 groups). Matching of these factors is important because age and peripheral hearing could influence EFR strengths (Anderson et al. 2012; Ananthakrishnan et al. 2016), whereas handedness is associated with hemispheric specialization (Carey and Johnstone 2014; Willems et al. 2014). The matching thus minimized possible confounds on any tDCS effects caused by these factors.

tDCS was applied at 1 mA for 25 min with the currents ramping up/down for 15 s at the stimulation onset/offset. For Sham, actual stimulation was applied for only 30 s in total (15 s ramping up and down respectively) at the very beginning of the 25-min period. This created the usual sensations associated with tDCS but without actual stimulation during the remainder of the run. The electrode configuration for each participant in the Sham group was randomly chosen from (1) to (4) except that it was ensured that the active electrode was positioned either on the left or on the right auditory cortex for an equal number of participants. Compared with other stimulation techniques like TMS that emits loud clicking sound during real stimulation, tDCS is silent, hence avoiding acoustic confounds on blinding participants. After the experiment, most participants, including those in the Sham group, orally reported that they believed they had received real stimulation. All experimental sessions were conducted during daytime (mornings or early afternoons) and all participants reported that they had at least 6 h sleep the night before which ensured adequate cortical plasticity triggered by tDCS (Salehinejad et al. 2019).

During tDCS, participants completed a pitch discrimination task while they listened to sound stimuli over a loudspeaker 1 meter in front of them in the same sound-shielded booth used for the EFR recordings. Three short complex tones (400 ms) calibrated at 75 dB SL at the 1-meter position were presented on each trial. The task was an "ABX" task. In each trial, 2 tones "A" and "B" with different F_0 (one of which had the same F_0 at 136 Hz as for the syllable used in the EFR recording) were played consecutively followed by a third tone "X" randomly selected from "A" or "B." Participants had to identify whether "X" was the same as "A" or "B." They gave their best guess if unsure of the answer. The process followed a "2-down, 1-up" adaptive procedure: F_0 difference between "A" and "B" decreased by $\sqrt{2}$ times following 2 consecutive correct trials and increased by $\sqrt{2}$ times following an incorrect trial. No feedback about response accuracy was provided. Half-minute breaks were taken every 4 min. The total number of completed trials was between 160 and 200 based on each participant's own pace. This task was included during tDCS because tDCS preferentially modulates neural networks that are currently active (Reato et al. 2010; Ranieri et al. 2012; Bikson and Rahman 2013). Concurrent tDCS and the pitch discrimination task could therefore maintain auditory cortical activity during neurostimulation, hence maximizing the effect of tDCS on neural excitability.

EEG Signal Processing

All EEG signal processing was conducted via Matlab R2017a (The Mathworks).

Preprocessing

As mentioned, the EFR was captured from Cz. The EEG signals were first re-referenced to the bilateral earlobes and bandpass filtered between 90 and 4000 Hz using a second-order zero-phase Butterworth filter. The filtered signals were then segmented and baseline-corrected (subtracting the average of the 50-ms prestimulus period) for each sweep. Sweeps that exceeded $\pm 25\,\mu$V were rejected to minimize movement artifacts. The resultant rejection rates were less than 2.5% averaged across participants in all cases (pre- and post-tDCS in the 5 stimulation groups for both left and right ear conditions).

EFR Magnitudes

FFRs with the positive and negative polarities (FFR_{Pos} and FFR_{Neg}) were first obtained by temporally averaging the pre-processed signals across sweeps with the respective polarities. EFR was measured as the response to the envelopes of F_0 and its harmonics by adding FFR_{Pos} and FFR_{Neg} (Aiken and Picton 2008). The addition of responses to the syllables with two polarities minimized the responses to TFS and cochlear microphonics, so that purer responses to envelopes were obtained to reflect the encoding of speech periodicity (Aiken and Picton 2008). A representative sample of EFR (waveform and spectrogram) is shown as Figure 2B. Here, EFR_{F0} and EFR_{2F0} (EFR at F_0 and its second harmonic, $2F_0$) that dominate the power of EFR were focused on. In contrast to higher harmonics (≥ 3) of EFR that may reflect distortion products resulting from nonlinear auditory response on the basilar membrane, EFR_{F0} and EFR_{2F0} reflect neural phase locking to sound periodicity in the central auditory systems (Smalt et al. 2012). Although it is expected that EFR_{F0} plays the major role in the phase locking, EFR_{2F0} also makes contributions (Aiken and Picton 2008) because of the nonsinusoidal characteristics of speech periodicity (Holmberg et al. 1988; also see discussions in Smalt et al. 2012). Magnitudes of EFR_{F0} and EFR_{2F0} were then measured following the procedure as follows:

1. Optimal neural lags relative to the syllable onset that generated the maximal stimulus-to-response correlations (Krizman and Kraus 2019) were first obtained for EFR. Specifically, both the syllable stimulus and EFR waveform were bandpass filtered at 126–146 and 262–282 Hz (i.e., 20 Hz bandwidth with center frequencies at F_0 and $2F_0$) using a second-order zero-phase Butterworth filter. 100 ms pre- and poststimulus periods were included during filtering to prevent filtering-induced boundary artifacts from contaminating the waveform at the stimulus period. Cross-correlations between the filtered stimulus and EFR were then conducted over a range of time delays (EFR lagged behind the stimulus) for the stimulus period at F_0 and $2F_0$, respectively. This range was set at 6–16 ms, for which 5–15 ms are latencies when EFR starts to occur in the brainstem (Chandrasekaran and Kraus 2010) with an additional 1 ms for sound transmission through the earphone plastic tube to the cochlea. The time delay that corresponded to the maximum absolute correlation value was treated as the optimal neural lag.

2. Magnitudes of EFR were then measured (for EFR_{F0} and EFR_{2F0} separately). The EFR waveform was windowed with a 5-ms rising and falling cosine ramp at the onset and offset. Note that the waveform was the one after preprocessing but prior to step 1, i.e., the bandpass filtering described in step 1 was for determining the optimal neural lags only, but not for measuring EFR magnitudes here in step 2. The window was set to lag behind the stimulus with the optimal neural lag obtained in the previous step (for EFR_{F0} and EFR_{2F0}, respectively). The windowed EFR waveform was then zero-padded to 1 s to allow for 1 Hz frequency resolution and the log-transformed

FFT-power spectrum [$10*\log_{10}(\text{power})$] was measured. EFR_{F0} and EFR_{2F0} magnitudes were taken as the powers centered at F_0 and $2F_0$ (averaged across 136 ± 2 Hz and 272 ± 2 Hz), respectively.

Statistical Analyses

All statistical analyses were conducted using SPSS Statistics 26.0 (IBM).

Baseline Characteristics

Before testing the aftereffects of tDCS, statistics were first conducted to check whether baseline (pre-tDCS) EFR characteristics were matched across stimulation groups. Linear mixed-effect regressions were conducted based on the restricted maximum likelihood approach. Baseline EFR magnitude and optimal neural lag were used as dependent variables. Stimulation Group (the 5 stimulation groups), Ear (left vs. right ear), and Harmonic (F_0 vs. $2F_0$) were fixed-effect factors, and Participant was the random-effect factor. The covariance matrix type was chosen among commonly used structures (first-order auto-regression, compound symmetry, diagonal, scaled identity, Toeplitz, and unstructured) to generate the lowest Bayesian Information Criterion values (i.e., best goodness of fit). The degrees of freedom were estimated via Satterthwaite approximation. Post hoc analyses were planned following significant interactions or main effects (Maxwell et al. 2017). Specifically, when there was a significant interaction, the whole dataset was split by the interactive factors and additional linear mixed-effect regressions were conducted at the respective levels within these factors. This procedure was repeated until a significant main effect was found. Pairwise comparisons were finally conducted between the underlying levels following a main effect (unless there were only 2 levels within the factor for the main effect, since in such cases the main effect already informed the significant difference between levels). P values for pairwise comparisons were corrected via FDR according to the multiple number of comparisons.

Aftereffects across Stimulation Groups and Ears

Analyses were then conducted to investigate how aftereffects of tDCS differed across stimulation groups and ears. Although we mainly focused on the aftereffects for EFR magnitude, we also tested aftereffects for the neural lag, because previous research also found that there is a potential association between the EFR neural lag and auditory cortical activities (Coffey et al. 2017b).

Linear mixed-effect regressions using the same fixed- (Stimulation Group, Ear, and Harmonic) and random-effect (participant) factors as for the baseline EFR were conducted. The dependent variables were the aftereffects of tDCS (i.e., differences in EFR magnitude and neural lag between post- and pre-tDCS). Here, we normalized the aftereffects to z values according to Sham (Grami et al. 2021) before the mixed-effect regressions were conducted. The z values were obtained as follows:

$$Z = \frac{X_{\text{tDCS}} - M_{\text{Sham}}}{\delta_{\text{Sham}}}$$

where X_{tDCS} denotes the raw aftereffects (either real tDCS or Sham), M_{Sham}, and δ_{Sham} denote the mean and standard deviation of the raw aftereffects across participants in the Sham group. Note that z values in the left and right ear conditions were

calculated separately, that is, z values for the left ear condition were calculated according to the Sham data obtained in the left ear condition, whereas z values for the right ear condition were calculated according to the Sham data obtained in the right ear condition. Compared with the raw values of the aftereffects, the z values account for the Sham effect that normalizes the data to better reflect aftereffects of the real tDCS (Grami et al. 2021).

Additional analyses were conducted using several potential confounding factors respectively as a fixed-effect covariate in the linear mixed-effect regressions. These factors include age, PTA (pre-tDCS), change in PTA (post- vs. pre-tDCS), and HI (all mean centered; missing data for the change in PTA were set as zero). Although all these factors were matched across stimulation groups (see "Experimental design"), such additional analyses were conducted to confirm that they did not contribute statistically to the aftereffects.

Laterality and Contralaterality of Aftereffects

The linear mixed-effect regressions described in the previous section tested the aftereffects across stimulation groups and ears. Importantly, they informed how aftereffects of real tDCS over the auditory cortex differed significantly from Sham. However, they were unable to directly test the main effect of the stimulated hemisphere (i.e., left vs. right auditory cortex) that reflects hemispheric laterality of cortical contributions and how laterality may depend on whether the listening ear was ipsilateral/contralateral to the stimulated site. This is because data of the Sham group were included in the analyses, and there was no definitive correspondence of Sham to specific stimulated hemispheres in the current design (see "Experimental Design").

Therefore, a further linear mixed-effect regression was conducted by excluding data in the Sham group to test the laterality and contralaterality of the aftereffects. Z value of the aftereffect was used as the dependent variable that had accounted for the Sham effect. We first conducted a preliminary analysis using the following 4 fixed-effect factors: Stimulated Hemisphere (the left vs. right auditory cortex), Contralaterality (the ipsilateral vs. contralateral ear to the stimulated hemisphere), Stimulation Type (anodal vs. cathodal), and Harmonic (F_0 vs. $2F_0$). Participant was the random-effect factor. We found that neither Stimulation Type nor Harmonic show significant main effects or interactions with other factors (all $P > 0.1$), indicating that effects of these 2 factors were negligible. Therefore, in order to reduce the analysis complexity and highlight the key effects, data of participants who received anodal and cathodal tDCS were grouped together, and the two harmonics were collapsed. A simplified linear mixed-effect regression was then conducted using only Stimulated Hemisphere and Contralaterality as fixed-effect factors and Participant as the random-effect factor. Post hoc analyses were conducted following a significant interaction effect (independent sample t-tests were conducted instead of testing main effects as there were only 2 levels for each fixed-effect factor). As results using all 4 fixed-effect factors in the preliminary analysis were similar to those in the simplified regression and led to the same conclusion, we only reported results for the simplified regression.

Also note that this linear mixed-effect regression was conducted only for EFR magnitude, but not for the neural lag. This is because significant aftereffects of real tDCS compared with Sham were only found for EFR magnitude (according to analyses in the previous section, see "Results"), following which laterality and contralaterality were further tested.

Timing Characteristics of Significant Aftereffects

Analyses were conducted to study when significant aftereffects on EFR magnitudes started to emerge by looking into variations of magnitudes across multiple time frames. Instead of measuring FFT power used for the linear mixed-effect regressions, we measured the power of temporal amplitude profile of the EFR waveform. This is because, empirically, the length of each time frame should be at least two F_0 cycles (14.7 ms) in order to accurately measure FFT power at F_0 (Liu and Morgan 2006). In contrast, a shorter length for each frame is needed for measuring power of amplitude profile that results in better temporal resolution. This is important because the emergence of subcortical/brainstem responses occur within a very narrow time window (5–15 ms after stimulus onset); hence, relatively fine-grained temporal resolution is needed to capture detailed magnitude variations so that emergence at subcortical and cortical responses can be disentangled.

Waveforms of EFR were first band-pass filtered using a second-order zero-phase Butterworth filter at 100–300 Hz (for F_0 plus $2F_0$) and 100–200 Hz (for F_0). The filtered waveforms were then Hilbert-transformed to obtain the Hilbert envelope as the temporal amplitude profile. The profile was then segmented into successive time frames time-locked to the stimulus onset. The length of each frame was set at 7.4 ms (corresponding to 1 cycle of F_0). We focused on the first 7 frames up till $\sim$50 ms (Frame 7 ranged at 44.4–51.8 ms after stimulus onset) that cover the latencies from brainstem to the primary auditory cortex (see Coffey et al. 2016) at which aftereffects could emerge. Magnitudes were measured as the log-power of the amplitude profile within each frame and were z-normalized according to Sham as conducted for the linear-mixed effect regressions.

The following two comparisons were then conducted via independent-sample *t*-tests for the magnitudes in each frame: (a) tDCS over the right auditory cortex (Right-AC Anodal and Cathodal) versus Sham in the left ear condition [Comparison (a)] and (b) tDCS over the right (Right-AC Anodal and Cathodal) versus left auditory cortex (Left-AC Anodal and Cathodal) in their respective contralateral ear conditions [Comparison (b)]. These comparisons were chosen because they were found to be significant according to results of the linear mixed-effect regressions (see "Results"). For simplicity, data for anodal and cathodal stimulation were grouped together for the *t*-tests, because no significant interactions involved with stimulation type (anodal vs. cathodal) were found in the previous linear mixed-effect regressions and significant aftereffects of Right-AC Anodal and Cathodal led to the same direction of changes (see "Results"). The comparisons were conducted via combining the 2 harmonics (F_0 plus $2F_0$) due to the lack of interactions involved with harmonic in the previous linear mixed-effect regressions. Comparisons were also conducted for the F_0 only because the upper frequency limit of cortical EFR is within the range of vocal pitch (<200 Hz, Guo et al. 2021). It is thus meaningful to look specifically into whether aftereffects at F_0 alone would start at a relatively later time frame at the cortical level. *P* values of the *t*-tests were all FDR-corrected according to the total number of steps (i.e., 7).

Pitch Discrimination Performances

Pitch discrimination was performed during the tDCS application (see "Experimental Design") and the change in discrimination performances was calculated for each participant. On each trial, pitch difference of the two complex tones ("A" and "B") was recorded. The initial and final discrimination thresholds were measured as the geometric means of the pitch differences across trials with the first 10 and final 10 reversals, respectively. The change in discrimination performances was then measured as the final threshold divided by the initial threshold.

We tested how changes in pitch discrimination differed across stimulation groups. A linear mixed-effect regression was conducted using the change in pitch discrimination performances as the dependent variable, Stimulation Group as the fixed-effect factor, and Participant as the random-effect factor. We also tested how changes in pitch discrimination were associated with changes in EFRs. We conducted Pearson's correlations between the changes in performances and aftereffects on EFR magnitudes in the respective stimulation groups. *P* values of correlations were FDR-corrected according to the total number of correlations conducted.

Results

Baseline Characteristics

Linear mixed-effect regressions were conducted for the baseline (pre-tDCS) EFR magnitude and neural lag.

For the EFR magnitude, there were significant main effects of Ear [$F(1, 116.140) = 12.978$, $P < 0.001$, greater magnitude in the left than the right ear with a small to medium effect size (Cohen's $d = 0.376$)] (Fig. 3) and Harmonic [$F(1, 90.631) = 67.029$, $P < 10^{-11}$, greater magnitude at F_0 than at $2F_0$ with a large effect size (Cohen's $d = 0.827$)]. There was no significant main effect of Stimulation Group or any 2- or 3-way interaction (all $P > 0.1$). Despite the lack of main effect of Stimulation Group, pairwise comparisons were still conducted in order to reassure that magnitudes were matched across stimulation groups. No significant differences were found between any 2 groups in each Ear and Harmonic condition [all $P > 0.1$, FDR-corrected according to multiple number of comparisons in each condition (i.e., 10)].

For the EFR neural lag, neither significant main effects nor any interactions were found (all $P > 0.1$). Similar to EFR magnitude, pairwise comparisons were still conducted for the neural lag despite the lack of main effect of Stimulation Group. No significant differences were found between any 2 groups in each Ear and Harmonic condition (all $P > 0.1$, FDR corrected).

Aftereffects across Stimulation Groups and Ears

Figure 4 shows the waveforms and FFT-power spectra of EFR pre- and post-tDCS in both ears, visualizing the aftereffects across stimulation groups and ears. Linear mixed-effect regressions were conducted for the z-normalized tDCS aftereffects on EFR. Statistical results are summarized in Table 1.

For the aftereffects on EFR magnitude, there were significant main effects of Stimulation Group [$F(4, 136.312) = 3.503$, $P = 0.009$] and ear [$F(1, 238.131) = 16.037$, $P < 10^{-4}$] and a significant (Stimulation Group $\times$ Ear) interaction [$F(1, 238.131) = 3.045$, $P = 0.018$]. Post hoc analyses following the significant interaction were conducted via splitting the dataset by Ear. Specifically, additional linear mixed-effect regressions were conducted in the left ear and right ear conditions, respectively. A significant main effect of Stimulation Group was found in the left [$F(4, 85) = 3.627$, $P = 0.009$] but not the right ear condition [$F(4, 123.058) = 0.919$, $P = 0.455$]. Pairwise comparisons were thus conducted between different stimulation groups using independent-sample *t*-tests in the left ear condition (collapsing the two harmonics due to no
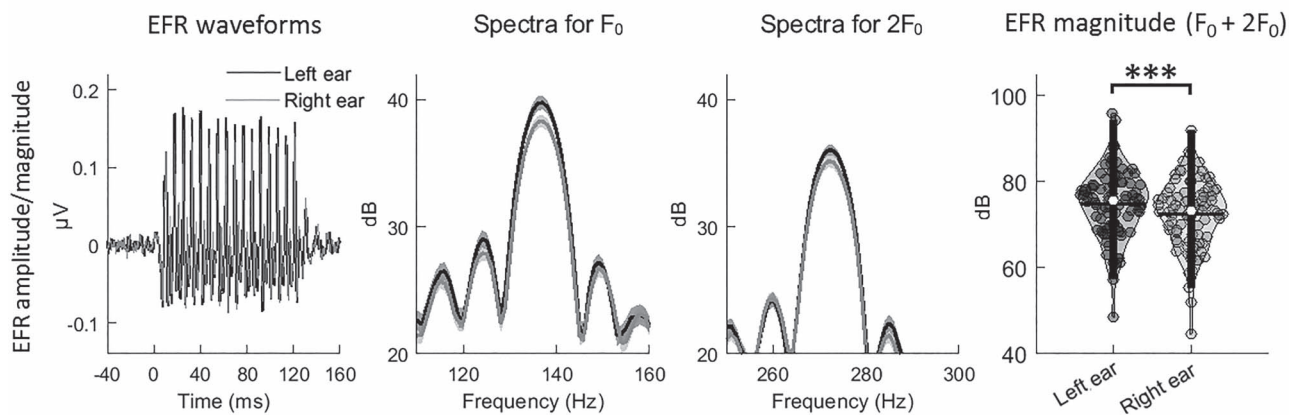
Figure 3. Comparison of baseline EFR magnitude between the left and the right ear conditions. From left to right shows the EFR waveforms, log-transformed FFT spectra for EFR at F_0 and $2F_0$, and the data distributions of EFR magnitudes. The EFR waveforms were the temporal grand-averaged waveforms across all participants. FFT spectra were obtained based on the windowed EFR (zero-padded to 1 second) time-locked to the optimal neural lags (for F_0 and $2F_0$, respectively; see the main text). The spectra show magnitudes peaking at 136 Hz (for F_0) and 272 Hz (for $2F_0$), respectively. Data distributions were illustrated as violin plots based on collapsing the two harmonics (addition of EFR magnitudes at F_0 and $2F_0$ for each participant due to the lack of significant interaction involving harmonic. In each violin plot, the white dot in the middle refers to the median value; the vertical and horizontal lines indicate the ± 1.5 interquartile range and the mean value, respectively. Dark and grey lines/dots denote data in the left and the right ear conditions, respectively. Shaded areas in the spectra cover the ranges of ± 1 standard error (SE) from the means. ***$P < 0.001$.

Table 1 Statistics results for linear mixed-effect regressions for the aftereffects (post- vs. pre-tDCS, z-normalized) on EFR magnitude and neural lag

| DV	Fixed-effect factors/interactions	df1	df2	F	P
Aftereffect on EFR magnitude (z-normalized)	Stimulation group × ear	4	238.131	3.045	**0.018***
	Stimulation group × Harmonic	4	238.131	0.472	0.756
	Ear × Harmonic	1	238.131	0.403	0.526
	Stimulation group × Ear × Harmonic	4	238.131	1.541	0.191
	Stimulation group	4	136.312	3.503	**0.009****
	Ear	1	238.131	16.037	**<10^{-4}*****
	Harmonic	1	238.131	1.499	0.222
Aftereffect on EFR neural lag (z-normalized)	Stimulation group × Ear	4	285.687	0.709	0.586
	Stimulation group × Harmonic	4	129.520	1.080	0.369
	Ear × Harmonic	1	295.884	7.887	**0.005****
	Stimulation group × Ear × Harmonic	4	295.884	1.016	0.399
	Stimulation group	4	140.489	0.861	0.489
	Ear	1	285.687	0.156	0.693
	Harmonic	1	129.520	1.397	0.239

Stimulation Group, Ear, and Harmonic were the fixed-effect factors and Participant was the random-effect factor. DV, df, F, and P denote dependent variables, degrees of freedom, F values, and P values, respectively. The degrees of freedom were estimated via Satterthwaite approximation. Significant P values (<0.05) are in bold.
*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

significant interactions involved with Harmonic). We found that Right-AC Anodal and Right-AC Cathodal resulted in significant decreases in EFR magnitude compared with Sham with large effect sizes [Right-AC Anodal vs. Sham: $t(34) = -3.100$, $P = 0.019$, Cohen's $d = 1.033$; Right-AC Cathodal vs. Sham: $t(34) = -3.351$, $P = 0.019$, Cohen's $d = 1.117$; P values were FDR corrected according to multiple number of comparisons (i.e., 10)]. In the right ear condition, pairwise comparisons between stimulation groups

were also conducted despite insignificant main effect of Stimulation Group to further reassure that aftereffects did not differ between groups. Indeed, no significant differences were found between any 2 groups (all $P > 0.5$, FDR corrected). Data distributions of the aftereffects on EFR magnitudes in the respectively ear conditions are illustrated in Figure 5.

For the aftereffects on EFR neural lag, neither significant main effects nor interactions were found, except for the
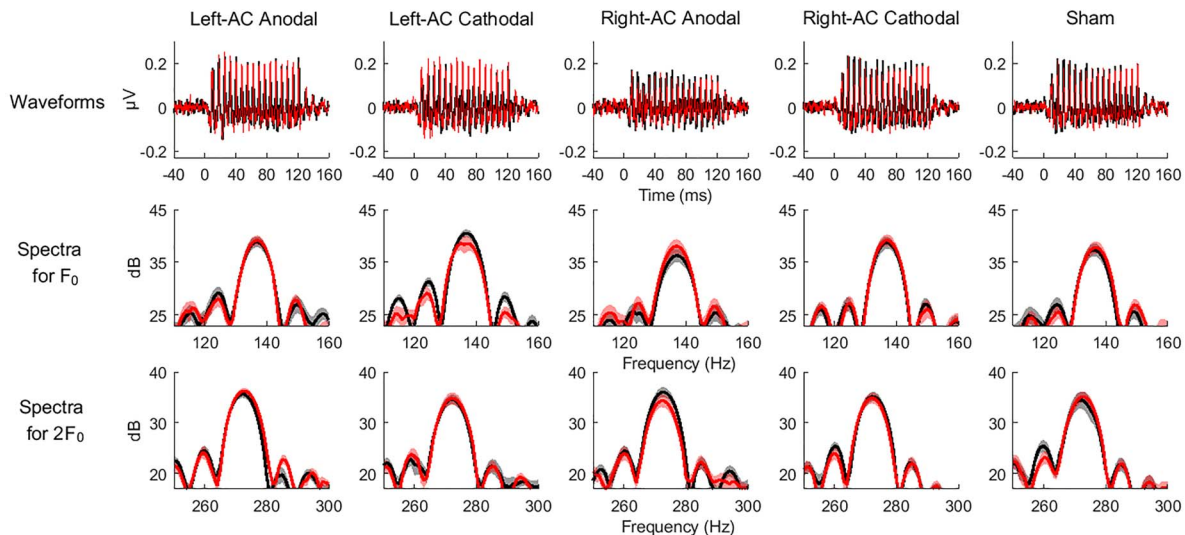
A. Left ear
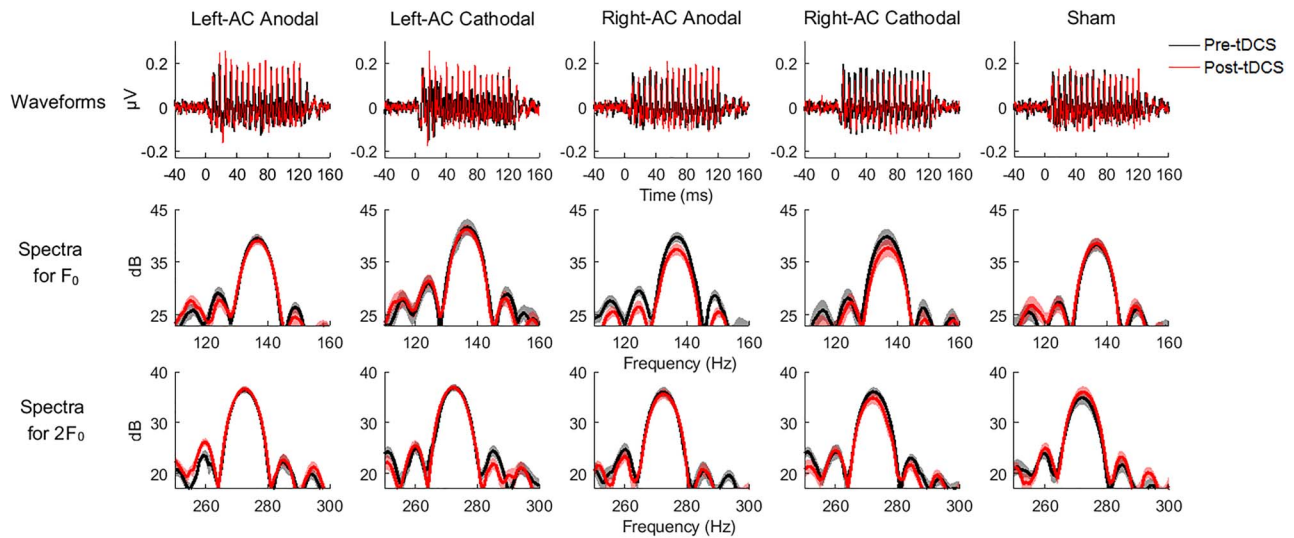
B. Right ear

Figure 4. Pre- and post-tDCS waveforms and power spectra for EFR. (A) and (B) show the waveforms and log-transformed FFT-power spectra in the left and right ear condition, respectively. Black and red lines indicate the pre- and post-tDCS data, respectively. From left to right shows different stimulation groups (Left-AC Anodal, Left-AC Cathodal, Right-AC Anodal, Right-AC Cathodal, and Sham). Upper panels show the grand-averaged EFR waveforms. Mid and lower panels show the FFT spectra based on the windowed EFR time-locked to the optimal neural lags for F_0 (mid) and $2F_0$ (lower). Shaded areas in the spectra cover the ranges of ± 1 SE from the means.

(Ear × Harmonic) interaction [$F(1, 295.884) = 7.887, P = 0.005$]. Post hoc analyses were thus required via splitting the dataset by Ear and Harmonic. Such effect, however, did not involve differences between stimulation groups (hence not within the current study's interest), the post hoc results are therefore not reported here. In addition, to further reassure that aftereffects on neural lag did not differ significantly between stimulation groups, pairwise comparisons were still conducted in the respective ear conditions despite insignificant main effects/interactions related to Stimulation Group. No significant differences were found between any stimulation groups in either ear (all $P > 0.4$, FDR corrected).

Furthermore, the same linear mixed-effect regressions were replicated by further including potential confounding factors

(age, PTA (pre-tDCS), change in PTA (post- vs. pre-tDCS), and HI) as fixed-effect covariates (including 1 confounder at a time). None of these factors changed the statistical significance of the original results and no significant main effects of them were found (all $P > 0.2$).

Laterality and Contralaterality of Aftereffects

A further linear mixed-effect regression was conducted for the aftereffects on EFR magnitude with data of the Sham group excluded. As mentioned in "Materials and Methods", neither stimulation type nor harmonics had main effects or interacted significantly with other factors. To reduce analysis complexity, data of anodal and cathodal were grouped together and the two
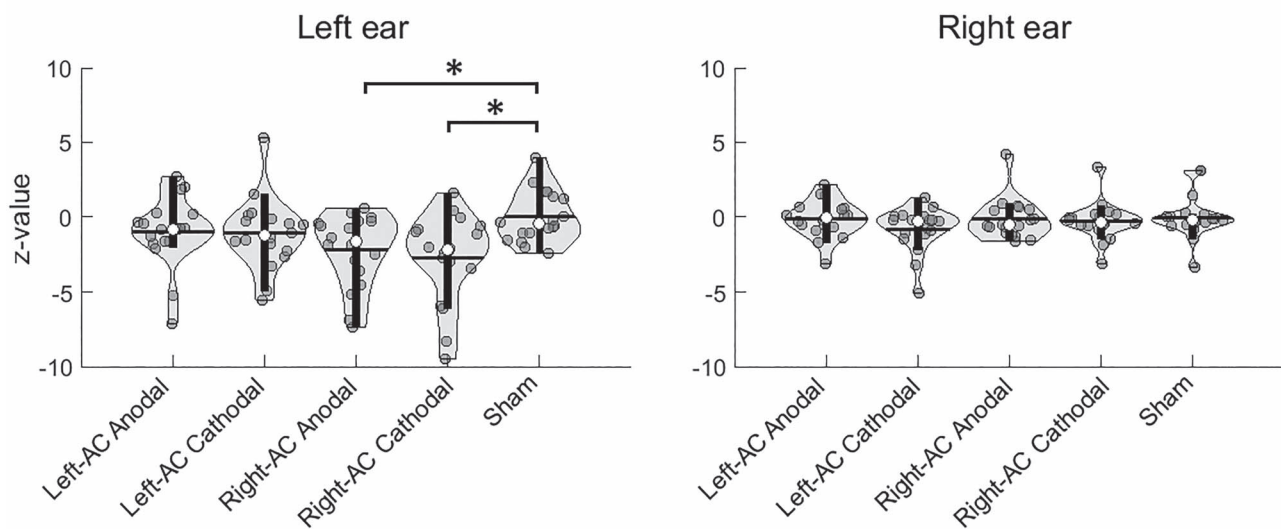
Figure 5. Data distributions of tDCS aftereffects on EFR magnitudes across stimulation groups. Aftereffects are shown as z values normalized according to Sham in the respective ear conditions. The distributions are illustrated as violin plots based on collapsing the two harmonics due to the lack of significant interaction involving harmonic. Data of different ear conditions are shown due to the significant (Stimulation Group × Ear) interaction. Post hoc analyses following the interaction showed a significant main effect of Stimulation Group in the left but not the right ear condition. Pairwise comparisons showed significant decreases in EFR magnitudes for the Right-AC Anodal and Right-AC Cathodal compared with Sham in the left ear condition. *$P < 0.05$, FDR-corrected according to multiple number of comparisons in each ear condition (i.e., 10).

Table 2 Statistics results for linear mixed-effect regressions for the aftereffects (post- vs. pre-tDCS, z-normalized) on EFR magnitude after excluding the Sham group

DV	Fixed-effect factors/interactions	df1	df2	F	P
Aftereffect on EFR magnitude (z-normalized)	Stimulated hemisphere × Contralaterality	1	70	18.394	**<10^{-4}*****
	Stimulated hemisphere	1	70	2.863	0.095
	Contralaterality	1	70	6.456	**0.013***

Stimulated Hemisphere and Contralaterality were the fixed-effect factors, and Participant was the random-effect factor. DV, df, F, and P denote dependent variables, degrees of freedom, F values, and P values, respectively. Significant P values (< 0.05) are in bold. *$P < 0.05$, ***$P < 0.001$.

harmonics were collapsed before the regression analysis. Data distributions of the aftereffects across stimulated hemispheres and the ipsilateral and contralateral ear conditions are illustrated in Figure 6. Statistical results of the regression are shown in Table 2.

A significant main effect of Contralaterality [$F(1, 70) = 6.456$, $P = 0.013$] and (Stimulated Hemisphere × Contralaterality) interaction [$F(1, 70) = 18.394$, $P < 10^{-4}$] were found. Post hoc pairwise comparisons were conducted following the significant interaction comparing: 1) aftereffects between different stimulated hemispheres (Left-AC vs. Right-AC tDCS) in the ipsilateral and contralateral ear conditions, respectively, and 2) aftereffects between ipsilateral and contralateral ear conditions for Left-AC and Right-AC tDCS, respectively. For 1), there was a significantly greater decrease in EFR magnitude for the Right-AC tDCS than the Left-AC tDCS in the contralateral ear condition ($t(52.421) = -3.948$, $P < 0.001$, unequal variance due to statistical significance of Levene's test; Cohen's $d = 0.913$), but not the ipsilateral ear condition. For 2), there was a significantly greater decrease in EFR magnitude in the contralateral than the

ipsilateral ear for the Right-AC tDCS ($t(50.456) = -4.514$, $P < 10^{-4}$, unequal variance due to statistical significance of Levene's test; Cohen's $d = 1.064$), but not Left-AC tDCS. Note that both significant effects had large effect sizes (Cohen's $d > 0.9$).

Besides the post hoc analyses described above, an additional pairwise comparison was conducted between Left-AC and Right-AC tDCS in the left ear (ipsilateral to the Left-AC tDCS but contralateral to the Right-AC tDCS) condition. This was to reassure whether the right-hemispheric laterality also occurred for the same ear when the ear was contralateral to the right auditory cortex. Indeed, we found significantly greater decrease in EFR magnitude for the Right-AC than the Left-AC tDCS in the left ear condition ($t(70) = 2.389$, $P = 0.020$, Cohen's $d = 0.563$). This significant effect is, however, not as strong as the laterality in the contralateral ear condition as shown previously. The reason may be because Left-AC tDCS could also partly affect the excitability in the right auditory cortex due to wide dispersion of currents generated by tDCS (Unal and Bikson 2018), which may in turn influence EFR to some extent in the left ear condition.
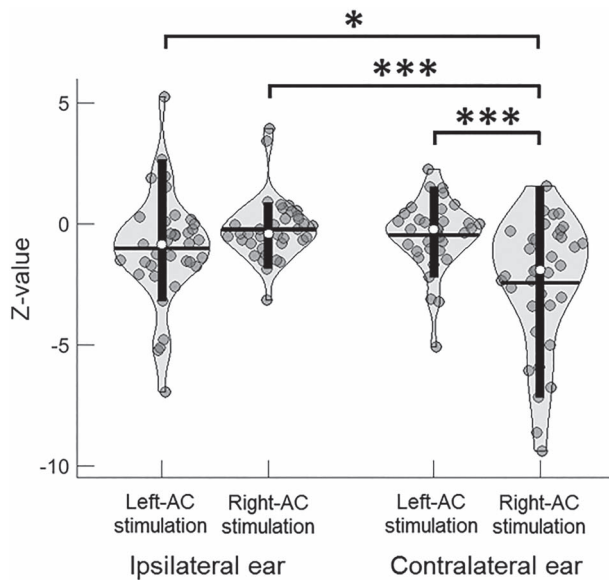
Figure 6. Data distributions of tDCS aftereffects on EFR magnitudes with data of Sham excluded. Aftereffects are shown as z values split by Stimulated Hemisphere (left/right auditory cortex) and Contralaterality (ears ipsilateral/contralateral to the stimulated hemispheres). As neither Stimulation Type (anodal/cathodal) nor Harmonic (F_0/$2F_0$) had main effects or interacted with other factors, data of anodal and cathodal were group together and the 2 harmonics were collapsed. Post hoc analyses following a significant (Stimulated Hemisphere × Contralaterality) interaction showed a significantly greater decrease in the EFR magnitude for tDCS over the right (Right-AC tDCS) than the left auditory cortex (Left-AC tDCS) in the contralateral but not ipsilateral ear condition, and a significantly greater decrease in magnitude in the contralateral than the ipsilateral ear when tDCS was applied over the right (Right-AC tDCS) but not left auditory cortex (Left-AC tDCS). In addition, Right-AC tDCS showed a significantly greater decrease in the EFR magnitude than the Left-AC tDCS in the left ear (ipsilateral to Left-AC but contralateral to Right-AC) condition. *$P < 0.05$, ***$P < 0.001$.

Timing Characteristics of Significant Aftereffects

Independent sample *t*-tests were conducted in different time frames comparing z values between the Right-AC tDCS and Sham in the left ear condition (Fig. 7A), and between the Right-AC and Left-AC tDCS in their respective contralateral ear conditions (Fig. 7B). The comparisons were conducted when combining the 2 harmonics (F_0 plus $2F_0$) (see upper panels of Fig. 7) and for F_0 only (see lower panels of Fig. 7). The results show that significant aftereffects started to emerge from Frame 2 (7.4–14.8 ms after stimulus onset) for all comparisons (also see Table 3 for the *t*-statistics). *P* values for the *t*-tests were FDR-corrected according to the total number of frames (i.e., 7). All significant effects had medium to large effect sizes (all Cohen's *d* were in the range of either 0.5–0.8 or > 0.8, see Table 3). The results thus indicate that emergence of significant aftereffects started at a relatively early stage (before 15 ms after stimulus onset, which is likely at the subcortical level, see Chandrasekaran and Kraus 2010) along the central auditory systems.

Pitch Discrimination

A linear mixed-effect regression was conducted for the change in pitch discrimination performances. The results showed no significant main effect of Stimulation Group [$F(4, 85) = 1.036$, $P = 0.393$] and no significant differences in the change in performances between any 2 Stimulation Groups (all $P > 0.4$, FDR corrected). Also, Pearson's correlations were conducted

between the changes in performances and the aftereffects on EFR magnitude (collapsing the two harmonics) in both ears in the respective stimulation groups. No significant correlations were found (all $P > 0.3$, FDR corrected).

Summary of Results

To summarize, our results showed as follows. First, we confirmed that EFR magnitude and neural lag were matched across stimulation groups at the baseline, so any aftereffect should not be attributed to baseline differences. We further found a relatively small (small to medium effect size) but significant left ear laterality for the baseline EFR magnitude. Second, we showed a significant (Stimulation Group × Ear) interaction for the aftereffects on EFR magnitude. Post hoc analyses found that tDCS over the right auditory cortex resulted in significant changes (decreases) in magnitude compared with Sham in the left ear (contralateral to the right auditory cortex) condition. On the other hand, no significant aftereffects were found for tDCS over the left auditory cortex compared with Sham. Third, we showed a significant (Stimulated Hemisphere × Contralateral) interaction and a hemispheric laterality in which the decrease in EFR magnitude in the contralateral ear was significantly greater when tDCS was applied over the right than the left auditory cortex. Finally, analyses on timing characteristics showed that emergence of the significant aftereffects and laterality mentioned in the second and the third points started from the relatively early, likely subcortical stage in the auditory systems.

The results therefore addressed our hypotheses that tDCS over the right auditory cortex causes changes in EFR along the contralateral auditory pathway and validate the causal relationship between the right auditory cortex and EFR.

Discussion

The current study used a combined tDCS and EEG approach to test for a causal contribution of auditory cortex to speech-evoked EFR in healthy right-handed participants. This approach can inform us that the cortical contribution found in the previous studies is not merely a result of specific source localization techniques or observed association between cortical activations and EFR, but that EFR is casually related to neuroplastic changes induced by neurostimulation in the auditory cortex. The left and right auditory cortices were neurostimulated in different groups of participants and the aftereffects of tDCS on the EFR were examined during monaural listening to a repeated speech syllable. Our results showed that tDCS over the right auditory cortex resulted in significant decrease in EFR magnitude compared with Sham as well as tDCS over the left auditory cortex (i.e., hemispheric laterality) when the listening ear was contralateral to the stimulated site. No significant aftereffects were found for tDCS over the left auditory cortex. The results thus agree with previous studies that have shown a close relationship between the right auditory cortex and EFR (Coffey et al. 2016, 2017a, 2017b; Hartmann and Weisz 2019; Ross et al. 2020) and provide the very first evidence for a causal relationship.

Causal Contributions of the Right Auditory Cortex to EFR and the Hemispheric Laterality

A recent effort was made to study how EFR is modulated by inhibiting excitability of the right auditory cortex via rTMS (López-Caballero et al. 2020). rTMS was applied over the right
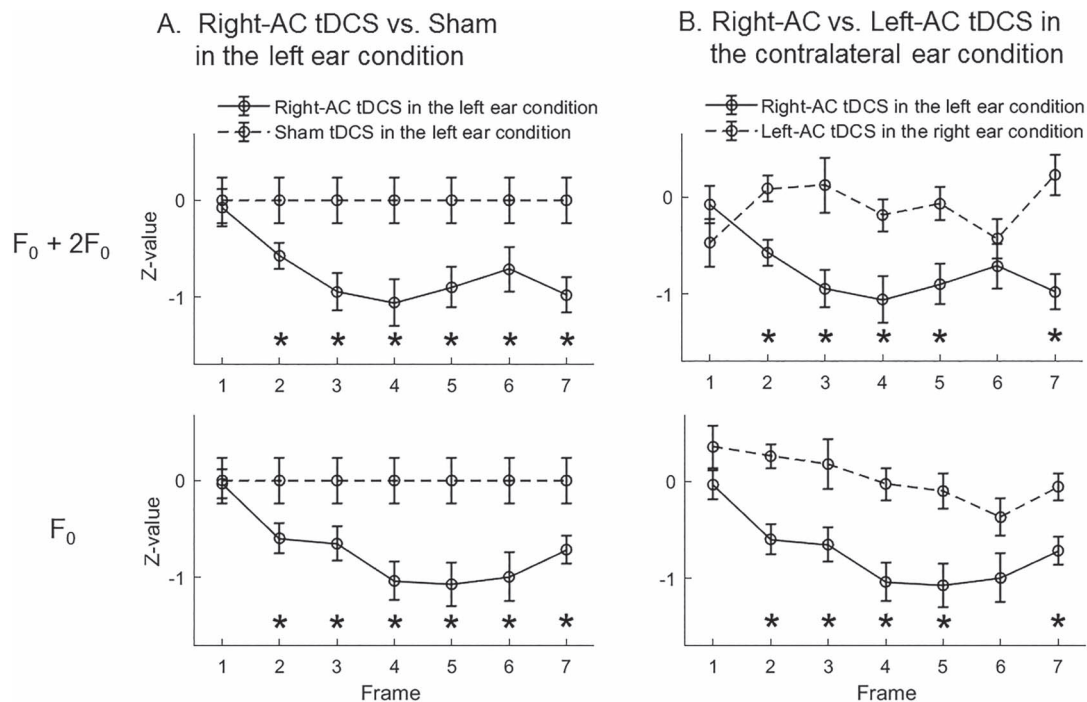
Figure 7. Timing characteristics of the aftereffects and hemispheric laterality. Z values are shown for the magnitudes of EFR amplitude profile that compare: (A) Right-AC tDCS with Sham in the left ear condition and (B) Right-AC with Left-AC tDCS in the respective contralateral ear conditions across time frames (Frame 1 to 7 time-locked to stimulus onset). The length of each frame was 7.4 ms corresponding to 1 cycle of F_0. Upper and lower panels respectively indicate the comparisons when collapsing the magnitudes at the 2 harmonics ($F_0 + 2F_0$) and at F_0 only. Error bars indicate the SEs. Asterisks indicate the significant differences for each frame via independent-sample t-tests. Note that the mean values of Sham were all zeros because they were z-normalized in each frame. *$P < 0.05$, FDR corrected according to multiple number of comparisons (i.e., 7).

Table 3 Comparisons of aftereffects on EFR magnitudes across time

			Frame 1 (0–7.4 ms)	Frame 2 (7.4–14.8 ms)	Frame 3 (14.8–22.2 ms)	Frame 4 (22.2–29.6 ms)	Frame 5 (29.6–37 ms)	Frame 6 (37–44.4 ms)	Frame 7 (44.4–51.8 ms)
Comparison (a)	$F_0 + 2F_0$	T	-0.236	−2.267	−2.948	−2.796	−2.652	−1.929	−3.200
		P	0.815	**0.028***	**0.005****	**0.007****	**0.011***	0.059	**0.002****
		d	0.068	0.654	0.851	0.807	0.766	0.557	0.924
	F_0	T	−0.121	−2.146	−2.173	−3.188	−2.988	−2.518	−2.760
		P	0.904	**0.037***	**0.034***	**0.002****	**0.004****	**0.015***	**0.008****
		d	0.035	0.620	0.627	0.920	0.863	0.727	0.797
Comparison (b)	$F_0 + 2F_0$	T	1.248	−3.460	−3.114	−2.993	−3.084	−0.930	−4.412
		P	0.216	**<0.001*****	**0.003****	**0.004****	**0.003****	0.355	**<10^{-4} *****
		d	0.294	0.816	0.734	0.705	0.727	0.219	1.040
	F_0	T	−1.464	−4.279	−2.689	−3.910	−3.347	−1.974	−3.352
		P	0.148	**<10^{-4} *****	**0.009****	**<0.001*****	**0.001****	0.052	**0.001****
		d	0.345	1.009	0.634	0.922	0.789	0.465	0.790

Comparisons were conducted between Right-AC tDCS and Sham in the left ear condition [comparison (a)] and between Right-AC tDCS and Left-AC tDCS in their respective contralateral ear condition [Comparison (b)]. Independent t-tests were conducted from Frame 1 to Frame 7 that combined the 2 harmonics ($F_0 + 2F_0$) and at F_0 only. The statistical results in the table correspond to the significant results illustrated in Figure 7 [Comparison (a) corresponds to Fig. 7A and Comparison (b) corresponds to Fig. 7B]. Numbers in the brackets denote the range for each time frame after the stimulus onset. The length of each frame was 7.4 ms (1 cycle of F_0). T, P, and d refer to the t values, P values, and Cohen's d, respectively. All P values were FDR corrected according to the number of comparisons (i.e., 7). Significant P values (<0.05) are in bold. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

auditory cortex and EFR magnitudes obtained via EEG were compared between post- and pre-TMS while participants binaurally listened to repeated syllables. No significant aftereffects on EFR were found. Here, we argue that the right auditory cortex may contribute to EFR mainly along the contralateral pathway. If

this is the case, the insignificant effect from the ipsilateral ear during binaural listening could conceal the genuine impact of the neurostimulation. We also argue that it is more appropriate to also include the condition where stimulation is applied over the left auditory cortex to test whether the aftereffect occurs

specifically in the right hemisphere and to test for hemispheric laterality. The present study, therefore, used an approach in which participants monaurally listened to stimuli to allow for testing contralaterality and neurostimulations were applied over both left and right auditory cortices so that hemispheric laterality were directly studied.

We found that tDCS over the right, but not left, auditory cortex resulted in significant changes in EFR magnitude compared with Sham in its contralateral (i.e., left) ear condition. Together with the hemispheric laterality found in the contralateral ear condition, our results argue for a causal role of the right auditory cortex in processing speech periodicity along the contralateral pathway in the central auditory systems. We suggest that the present study advances our understanding of the relationship between EFR and pitch processing in the right auditory cortex. Previous studies have shown that EFR is closely related to pitch perception. EFR strength can be enhanced by both short-term perceptual training of pitch discrimination (Carcagno and Plack 2011) as well as long-term musical experience (Musacchia et al. 2007; Wong et al. 2007; Strait et al. 2009; Bidelman et al. 2011). Furthermore, EFR has been used as an index of neural fidelity of linguistic pitch and the fidelity is greater in tonal language than in nontonal language speakers (Krishnan et al. 2004, 2005, 2009). Despite this, rather than reflecting the result of pitch extraction, EFR has been suggested to reflect subcortical responses to monaural temporal information (e.g., periodicity cues) that are important for extracting pitch of complex sounds (i.e., "pitch-bearing" information; Gockel et al. 2011). On the other hand, the process of pitch extraction itself takes place in the auditory cortex (Penagos et al. 2004; Bendor and Wang 2005; Puschmann et al. 2010) with a right hemispheric specialization (Zatorre and Belin 2001; Patterson et al. 2002; Hyde et al. 2008; Mathys et al. 2010; Albouy et al. 2013; Matsushita et al. 2015, 2021). In this respect, the current aftereffects of tDCS may reflect a top-down corticofugal modulation process in which the right auditory cortex affects the processing of pitch-bearing information that occurs at the subcortical level. A model of top-down modulation between auditory cortex and subcortex that controls the temporal dynamics of pitch processing has been proposed and stresses that this process is key for pitch perception (Balaguer-Ballester et al. 2009). Alternatively, although EEG mainly captures EFR signals originating from the brainstem (Bidelman 2015, 2018), cortical sources have been found dominated in the right hemisphere (Coffey et al. 2016, 2017a). It may be that the aftereffects reflected the changes in the cortical EFR activities. Therefore, additional analyses for the timing characteristics of when significant aftereffects started to emerge were further investigated in the current study. The results seem to support the proposal arguing for the top-down role of the auditory cortex. This will be further discussed in the next section "Timing characteristics of the significant aftereffects".

It is also noteworthy that we showed a laterality effect at the baseline as well, where EFR had significantly greater magnitude in the left than the right ear condition. This echoes previous findings showing right-hemispheric laterality for auditory steady-state responses (ASSRs) (but at lower frequencies of 40 and 80 Hz, Vanvooren et al, 2015; Ross et al. 2005; Luke et al. 2017). Both ASSRs and EFR are phase-locked responses to envelope modulations, implying that phase-locked envelope responses in general might be more prominent along the contralateral pathway from the left ear to the right auditory cortex. However, it is not clear whether and how auditory cortex contributes to this laterality for EFR. As such, the current results provide confirmatory evidence for a causal cortical contribution to this laterality. It would be interesting to see whether such causality could be replicated for ASSR in the future. It should also be noted that the present study was conducted specifically in right-handed participants. Although hemispheric laterality is associated with handedness (Carey and Johnstone 2014; Willems et al. 2014), the current results may not necessarily apply to people who are left-handed or ambidextrous.

Besides measuring EFR, the present study also used a pitch discrimination task during tDCS application so that the auditory cortex was kept active and the tDCS effects were optimized (Reato et al. 2010; Ranieri et al. 2012; Bikson and Rahman 2013). Unlike EFR, changes in pitch discrimination performances did not differ between stimulation groups and no correlations were found between changes in performances and EFR magnitudes in any group. This may be because, with our electrode configuration (active electrode on the auditory cortex and reference on the forehead above the contralateral eyebrow), currents generated by tDCS would pass through not only auditory cortices but also various brain regions due to its marked diffusive nature (Faria et al. 2011; Bai et al. 2014; Unal and Bikson 2018). Pitch discrimination involves not only sensory regions (i.e., auditory cortices) but also higher-order regions such as frontal and prefrontal cortices (Zatorre et al. 1992; Palomar-Garcia et al. 2020) and anterior cingulate cortex (Schwenzer and Mathiak 2011) that may have been stimulated by tDCS. We thus argue that it is not surprising that no significant results were found for pitch discrimination because of such diffusive effects of tDCS.

Another important issue that needs discussion is that the present study focused on EFR that reflects neural encoding of periodicity envelope at F_0 and its harmonic ($2F_0$). Besides periodicity envelope, TFS, particularly at the resolved harmonic region, is also essential for pitch perception (Moore 2014). Here, we did not include responses to TFS because they mainly reflect encoding of sounds at the auditory periphery or cochlear microphonics (Aiken and Picton 2008) rather than responses at the central auditory systems. Also, there may be electrical artifacts when recording responses to TFS, whereas EFR was obtained by adding responses to the positive and negative polarities such that artifacts are minimized (Skoe and Kraus 2010). It is therefore valuable for the future to explore an artifact-free response to TFS located in the central auditory systems and how it may relate to cortical hemispheric laterality. In contrast to TFS, periodicity envelope is most important for pitch perception when auditory signals are dominated by unresolved harmonics (Oxenham et al. 2009; Moore and Gockel 2011). Furthermore, periodicity encoding is more important when TFS is not well perceived or available in clinical populations. For example, evidence showed that encoding of periodicity envelope is less deteriorated than TFS by aging and hearing loss (Halliday et al. 2019; Moore 2021). In hearing protheses such as cochlear implants (CI), periodicity envelope rather than TFS is effectively conveyed, so CI recipients would rely mainly on periodicity encoding to perceive pitch (Hofmann and Wouters 2010; Wouters et al. 2015; Gransier et al. 2020). Therefore, it would be also meaningful for future studies to investigate cortical contributions to EFR in hearing-impaired populations.

Timing Characteristics of the Significant Aftereffects

Analyses of the timing characteristics showed that the significant aftereffects and hemispheric laterality observed in the

present study emerged at a relative early, subcortical time window (~7–15 ms after stimulus onset). This early emergence happened both when collapsing EFR magnitudes at F_0 and $2F_0$ and for magnitudes at F_0 only.

Therefore, such results indicate that the aftereffects may be introduced via a top-down corticofugal modulation process in which excitability changes in the right auditory cortex affected EFR that occurs at the subcortical level. It is noteworthy that "cortical contribution" we refer to here is not only limited to possible contributions of cortical sources of EFR, but also corticofugal modulation that involves changes in the right auditory cortex. Another argument may be that, although tDCS was applied over the auditory cortices, the aftereffects could be due to the impact of tDCS directly on the brainstem without cortical contributions. However, we argue that this latter proposal was not likely to be the case, because if it were, impacts imposed by tDCS on the brainstem are expected to be similar between stimulations over different hemispheres. In such case, the laterality of aftereffects should be determined only by factors below the cortical level (i.e., ipsilateral/contralateral ear) regardless of which hemispheres were stimulated. Our current results, however, showed a significant interaction between Stimulated Hemisphere and Contralaterality (ear ipsilateral/contralateral to the stimulated hemisphere). Hemispheric laterality was shown in which tDCS resulted in significantly greater aftereffects over the right than the left auditory cortex in the contralateral, but not ipsilateral, ear condition. This therefore argues against the possibility that the aftereffects were consequences of the tDCS directly imposed on the brainstem and argues for the top-down corticofugal modulation process.

The auditory corticofugal modulation has been well established in animal studies (Suga et al. 2000; Bajo et al. 2010; Bajo and King 2013; Suga 2020). Specifically, excitatory electrical stimulations in the auditory cortex can facilitate the responses of IC neurons that have the same tuning frequencies as the stimulated cortical neurons (Bajo and King 2013; Suga 2020). On the other hand, inactivating auditory cortex reduces the responses of IC neurons (Zhang and Suga 1997; Yan and Suga 1999; Bajo and King 2013). Therefore, a possible mechanism that underlies the current observed aftereffects would be that tDCS altered the excitability in the auditory cortex. This would then lead to hyperpolarization of subcortical postsynaptic neurons through the efferent connections from the auditory cortex to the subcortex. The hyperpolarization would raise the thresholds for the occurrence of action potentials in response to speech periodicity, reflected by the decrease in EFR magnitude. The interesting result we observed here is that this process happens specifically for the right auditory cortex along the contralateral pathway. The human auditory cortex has been shown to have sharper frequency tuning in the right than the left hemisphere (Liégeois-Chauvel et al. 1999, 2001), suggesting greater number of neurons responsible for spectral analyses of periodicity information that support pitch perception in the right hemisphere (Zatorre et al. 2002). Therefore, top-down corticofugal modulation for pitch perception (Balaguer-Ballester et al. 2009) may be prominent in the right hemisphere and hence more susceptible to external neurostimulations compared with Sham and stimulations over the left hemisphere.

Despite our observations of the timing characteristics, some limitations are yet to be addressed. First, previous research argued that cortical EFRs have upper frequency limit at ~ 100 Hz (Bidelman 2018). However, recent data have shown that cortical EFRs also include frequencies that cover a more common range of human vocal pitch (100–140 Hz, Ross et al. 2020; 100–200 Hz, Guo et al. 2021). The present study used a speech syllable with F_0 at 136 Hz that falls within the 100–200 Hz range. Nonetheless, we had not obtained compelling results here in order to conclude that the aftereffects were directly from the cortex. Future studies may use a lower F_0 to clarify whether the timing characteristics are different across F_0 frequencies. Second, using scalp-recorded EEG, we were not able to spatially localize the aftereffects. Therefore, an even better approach for future studies to consider is to use technique like MEG that is capable of localizing the aftereffects at the subcortical and/or cortical level (Coffey et al. 2016). Combining the timing and spatial information could provide a more comprehensive image showing us both when and where the effects start to emerge.

Neurophysiological Consequences of tDCS

An intriguing finding of the current study is that anodal and cathodal tDCS resulted in the same direction of changes, both causing decreases in EFR magnitude for tDCS over the right auditory cortex in the contralateral (i.e., left) ear condition. Conventionally, anodal and cathodal stimulations reflect depolarization and hyperpolarization of neurons, respectively, which should lead to opposite directions of aftereffects (Jacobson et al. 2012). However, it is not unusual that tDCS has polarity-independent effects due to the underlying complexity of its neurophysiological consequences. For example, several studies have shown that anodal and cathodal tDCS have the same effects on excitability of motor cortex (Antal et al. 2007), motor learning (de Xivry et al. 2011), cerebellar functions for working memory (Ferrucci et al. 2008), and visuomotor learning (Shah et al. 2013). The mechanisms underlying excitatory/inhibitory consequences of anodal and cathodal stimulations may be the changes in concentrations of relevant neurotransmitters. Stagg et al. (2009) found that anodal tDCS over the motor cortex causes decreases in gamma-amino butyric acid (GABA) concentration that lead to excitation, whereas cathodal tDCS also causes decreases in GABA but greater concurrent decreases in glutamate that lead to cortical inhibition. A recent study applied tDCS over the human auditory cortex and found that both anodal and cathodal stimulations resulted in increases in relative concentration of GABA to glutamate (Heimrath et al. 2020). This is consistent with the current results that both anodal and cathodal tDCS led to decreased EFR magnitude, indicating that both stimulation types may have introduced neural inhibition in the auditory cortex. It would be worthwhile for future studies to investigate how changes in concentrations of neurotransmitters by neurostimulation relate to changes in EFR, which can help us better understand the underlying mechanisms of cortical contributions to EFR.

Speech-Specific or Domain-General

Our auditory stimulus used to obtain EFR was a speech syllable. Another important question is whether our results are speech-specific or could be generalized to nonspeech stimuli. It has been shown that the right auditory cortical contributions are similar when using a speech syllable and a music note (Coffey et al. 2017a). However, speech and music notes share a great deal of physical properties (e.g., spectral complexity and harmonic structures), which may lead to such similarity. It is not clear whether using other stimuli such as complex tones, amplitude-modulated tones or iterated rippled noise (Krishnan et al. 2009; Ananthakrishnan and Krishnan 2018) would lead to the same or

different effects as observed here. A future direction is then to examine whether the causal contribution of the right auditory cortex is speech-specific or domain-general.

Conclusion

We showed that tDCS over the right, but not left, auditory cortex resulted in significant changes in speech-evoked EFR magnitude compared with Sham in the contralateral (i.e., left) ear condition. Crucially, we also showed a hemispheric laterality in which the aftereffect was greater when tDCS was applied over the right than the left auditory cortex in the contralateral ear condition. Furthermore, we showed that the aftereffects and the laterality emerged from the relatively early, subcortical stages, indicating a top-down corticofugal modulation of the right auditory cortex on the subcortex. The current results thus validate the previous findings that the right auditory cortex makes significant contributions to EFR by establishing a causal relationship between the two. To our knowledge, this is the first evidence for this causality. Our findings should advance the understanding of how periodicity and pitch information are processed along the central auditory pathways in the human brain.

Funding

Notes

References

Aiken SJ, Picton TW. 2008. Envelope and spectral frequency-following responses to vowel sounds. *Hear Res.* 245:35–47.

Albouy P, Mattout J, Bouet R, Maby E, Sanchez G, Aguera PE, Daligault S, Delpuech C, Bertrand O, Caclin A, et al. 2013. Impaired pitch perception and memory in congenital amusia: the deficit starts in the auditory cortex. *Brain.* 136(5):1639–1661.

Ananthakrishnan S, Krishnan A. 2018. Human frequency following responses to iterated rippled noise with positive and negative gain: differential sensitivity to waveform envelope and temporal fine-structure. *Hear Res.* 367:113–123.

Ananthakrishnan S, Krishnan A, Bartlett E. 2016. Human frequency following response: neural representation of envelope and temporal fine structure in listeners with normal hearing and sensorineural hearing loss. *Ear Hear.* 37(2):e91.

Anderson S, Parbery-Clark A, Yi HG, Kraus N. 2011. A neural basis of speech-in-noise perception in older adults. *Ear Hear.* 32(6):750.

Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. 2012. Aging affects neural precision of speech encoding. *J Neurosci.* 32(41):14156–14164.

Antal A, Terney D, Poreisz C, Paulus W. 2007. Towards unravelling task-related modulations of neuroplastic changes induced in the human motor cortex. *Eur J Neurosci.* 26(9):2687–2691.

Bai S, Dokos S, Ho KA, Loo C. 2014. A computational modelling study of transcranial direct current stimulation montages used in depression. *Neuroimage.* 87:332–344.

Bajo VM, King AJ. 2013. Cortical modulation of auditory processing in the midbrain. *Frontiers in Neural Circuits.* 6:114.

Bajo VM, Nodal FR, Moore DR, King AJ. 2010. The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat Neurosci.* 13:253–260.

Balaguer-Ballester E, Clark NR, Coath M, Krumbholz K, Denham SL. 2009. Understanding pitch perception as a hierarchical process with top-down modulation. *PLoS Comput Biol.* 5(3):e1000301.

Banai K, Abrams D, Kraus N. 2007. Sensory-based learning disability: insights from brainstem processing of speech sounds. *Int J Audiol.* 46(9):524–532.

Bendor D, Wang X. 2005. The neuronal representation of pitch in primate auditory cortex. *Nature.* 436(7054):1161–1165.

Bidelman GM, Krishnan A, Gandour JT. 2011. Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch. *Eur J Neurosci.* 33(3):530–538.

Bidelman GM. 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 GM. 2018. Subcortical sources dominate the neuro-electric auditory frequency-following response to speech. *Neuroimage.* 175:56–69.

Bidelman GM, Lowther JE, Tak SH, Alain C. 2017. Mild cognitive impairment is characterized by deficient brainstem and cortical representations of speech. *J Neurosci.* 37(13): 3610–3620.

Bikson M, Rahman A. 2013. Origins of specificity during tDCS: anatomical, activity-selective, and input-bias mechanisms. *Front Hum Neurosci.* 7:688.

Boroda E, Sponheim SR, Fiecas M, Lim KO. 2020. Transcranial direct current stimulation (tDCS) elicits stimulus-specific enhancement of cortical plasticity. *Neuroimage.* 211:116598.

Carcagno S, Plack CJ. 2011. Subcortical plasticity following perceptual learning in a pitch discrimination task. *J Assoc Res Otolaryngol.* 12(1):89–100.

Carey DP, Johnstone LT. 2014. Quantifying cerebral asymmetries for language in dextrals and adextrals with random-effects meta analysis. *Front Psychol.* 5:1128.

Cha K, Zatorre RJ, Schönwiesner M. 2016. Frequency selectivity of voxel-by-voxel functional connectivity in human auditory cortex. *Cereb Cortex.* 26(1):211–224.

Chandrasekaran B, Kraus N. 2010. The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology.* 47(2):236–246.

Coffey EB, Herholz SC, Chepesiuk AM, Baillet S, Zatorre RJ. 2016. Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat Commun.* 7(1):1–11.

Coffey EB, Musacchia G, Zatorre RJ. 2017a. Cortical correlates of the auditory frequency-following and onset responses: EEG and fMRI evidence. *J Neurosci.* 37(4):830–838.

Coffey EB, Chepesiuk AM, Herholz SC, Baillet S, Zatorre RJ. 2017b. Neural correlates of early sound encoding and their relationship to speech-in-noise perception. *Front Neurosci.* 11:479.

Coffey EB, Nicol T, White-Schwoch T, Chandrasekaran B, Krizman J, Skoe E, Zotorre RJ, Kraus N. 2019. Evolving perspectives on the sources of the frequency-following response. *Nat Commun.* 10(1):1–10.

Cunningham J, Nicol T, Zecker SG, Bradlow A, Kraus N. 2001. Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. *Clin Neurophysiol*. 112(5):758–767.

de Xivry JJ, Marko MK, Pekny SE, Pastor D, Izawa J, Celnik P, Shadmehr R. 2011. Stimulation of the human motor cortex alters generalization patterns of motor learning. *J Neurosci*. 31(19):7102–7110.

Encina-Llamas G, Harte JM, Dau T, Shinn-Cunningham B, Epp B. 2019. Investigating the effect of cochlear synaptopathy on envelope following responses using a model of the auditory nerve. *J Assoc Res Otolaryngol*. 20(4):363–382.

Faria P, Hallett M, Miranda PC. 2011. A finite element analysis of the effect of electrode area and inter-electrode distance on the spatial distribution of the current density in tDCS. *J Neural Eng*. 8(6):066017.

Ferrucci R, Marceglia S, Vergari M, Cogiamanian F, Mrakic-Sposta S, Mameli F, Zago S, Babieri S, Priori A. 2008. Cerebellar transcranial direct current stimulation impairs the practice-dependent proficiency increase in working memory. *J Cogn Neurosci*. 20(9):1687–1697.

Fujihira H, Shiraishi K. 2015. Correlations between word intelligibility under reverberation and speech auditory brainstem responses in elderly listeners. *Clin Neurophysiol*. 126(1):96–102.

Gockel HE, Carlyon RP, Mehta A, Plack CJ. 2011. The frequency following response (FFR) may reflect pitch-bearing information but is not a direct representation of pitch. *J Assoc Res Otolaryngol*. 12(6):767–782.

Gorina-Careta N, Kurkela JL, Hamalainen 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*. 231:117866.

Grami F, De Marco G, Bodranghien F, Manto M, Habas C. 2021. Cerebellar transcranial direct current stimulation reconfigurates static and dynamic functional connectivity of the resting-state networks. *Cerebellum & Ataxias*. 8(1):1–2.

Gransier R, Carlyon RP, Wouters J. 2020. Electrophysiological assessment of temporal envelope processing in cochlear implant users. *Sci Rep*. 10(1):1–14.

Guo N, Si X, Zhang Y, Ding Y, Zhou W, Zhang D, Hong B. 2021. Speech frequency-following response in human auditory cortex is more than a simple tracking. *Neuroimage*. 226:117545.

Halliday LF, Rosen S, Tuomainen O, Calcus A. 2019. Impaired frequency selectivity and sensitivity to temporal fine structure, but not envelope cues, in children with mild-to-moderate sensorineural hearing loss. *J Acoust Soc Am*. 146(6):4299–4314.

Hartmann T, Weisz N. 2019. Auditory cortical generators of the frequency following response are modulated by intermodal attention. *Neuroimage*. 203:116185.

Heimrath K, Brechmann A, Blobel-Lüer R, Stadler J, Budinger E, Zaehle T. 2020. Transcranial direct current stimulation (tDCS) over the auditory cortex modulates GABA and glutamate: a 7 T MR-spectroscopy study. *Sci Rep*. 10(1):1–8.

Hofmann M, Wouters J. 2010. Electrically evoked auditory steady state responses in cochlear implant users. *J Assoc Res Otolaryngol*. 11(2):267–282.

Holmberg EB, Hillman RE, Perkell JS. 1988. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J Acoust Soc Am*. 84(2):511–529.

Hornickel J, Kraus N. 2013. Unstable representation of sound: a biological marker of dyslexia. *J Neurosci*. 33(8):3500–3504.

Hyde KL, Peretz I, Zatorre RJ. 2008. Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia*. 46(2):632–639.

Impey D, Knott V. 2015. Effect of transcranial direct current stimulation (tDCS) on MMN-indexed auditory discrimination: a pilot study. *J Neural Transm*. 122:1175–1185.

Jacobson L, Koslowsky M, Lavidor M. 2012. tDCS polarity effects in motor and cognitive domains: a meta-analytical review. *Exp Brain Res*. 216(1):1–10.

Krishnan A, Xu Y, Gandour JT, Cariani PA. 2004. Human frequency-following response: representation of pitch contours in Chinese tones. *Hear Res*. 189:1–12.

Krishnan A, Xu Y, Gandour J, Cariani P. 2005. Encoding of pitch in the human brainstem is sensitive to language experience. *Cogn Brain Res*. 25(1):161–168.

Krishnan A, Gandour JT, Bidelman GM, Swaminathan J. 2009. Experience dependent neural representation of dynamic pitch in the brainstem. *Neuroreport*. 20(4):408.

Krizman J, Kraus N. 2019. Analyzing the FFR: A tutorial for decoding the richness of auditory function. *Hear Res*. 382:107779.

Lehmann A, Schönwiesner M. 2014. Selective attention modulates human auditory brainstem responses: relative contributions of frequency and spatial cues. *PLoS One*. 9(1):e85442.

Liegeois-Chauvel C, De Graaf JB, Laguitton V, Chauvel P. 1999. Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb Cortex*. 9(5): 484–496.

Liegeois-Chauvel C, Giraud K, Badier JM, 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(1):117–132.

Liu JG, Morgan GL. 2006. FFT selective and adaptive filtering for removal of systematic noise in ETM+ imageodesy images. *IEEE Trans Geosci Remote Sens*. 44(12):3716–3724.

López-Caballero F, Martin-Trias P, Ribas-Prats T, Gorina-Careta N, Bartrés-Faz D, Escera C. 2020. Effects of cTBS on the frequency-following response and other auditory evoked potentials. *Front Hum Neurosci*. 14:250.

Luke R, De Vos A, Wouters J. 2017. Source analysis of auditory steady-state responses in acoustic and electric hearing. *Neuroimage*. 147:568–576.

Mai G, Tuomainen J, Howell P. 2018. Relationship between speech-evoked neural responses and perception of speech in noise in older adults. *J Acoust Soc Am*. 143(3):1333–1345.

Mai G, Schoof T, Howell P. 2019. Modulation of phase-locked neural responses to speech during different arousal states is age-dependent. *Neuroimage*. 189:734–744.

Mathys C, Loui P, Zheng X, Schlaug G. 2010. Non-invasive brain stimulation applied to Heschl's gyrus modulates pitch discrimination. *Front Psychol*. 1:193.

Matsushita R, Andoh J, Zatorre RJ. 2015. Polarity-specific transcranial direct current stimulation disrupts auditory pitch learning. *Front Neurosci*. 9:174.

Matsushita R, Puschmann S, Baillet S, Zatorre RJ. 2021. Inhibitory effect of tDCS on auditory evoked response: simultaneous MEG-tDCS reveals causal role of right auditory cortex in pitch learning. *Neuroimage*. 233:117915.

Maxwell SE, Delaney HD, Kelley K. 2017. *Designing experiments and analyzing data: a model comparison perspective*. 3rd ed. New York (US): Routledge.

Moore BC. 2014. *Auditory processing of temporal fine structure: Effects of age and hearing loss*. Singapore: World Scientific.

Moore BC. 2021. Effects of hearing loss and age on the binaural processing of temporal envelope and temporal fine structure information. *Hear Res*. 402:107991.

Moore BC, Gockel HE. 2011. Resolvability of components in complex tones and implications for theories of pitch perception. *Hear Res*. 276:88–97.

Musacchia G, Sams M, Skoe E, Kraus N. 2007. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc Natl Acad Sci*. 104(40):15894–15898.

Nitsche M, Paulus W. 2001. Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. *Neurology*. 57(10):1899–1901.

Oldfield RC. 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*. 9(1):97–113.

Oxenham AJ, Micheyl C, Keebler MV. 2009. Can temporal fine structure represent the fundamental frequency of unresolved harmonics? *J Acoust Soc Am*. 125(4):2189–2199.

Palomar-Garcia MA, Hernandez M, Olcina G, Adrian-Ventura J, Costumero V, Miro-Padilla A, Villar-Rodriguez E, Avila C. 2020. Auditory and frontal anatomic correlates of pitch discrimination in musicians, non-musicians, and children without musical training. *Brain Struct Funct*. 225(9):2735–2744.

Parbery-Clark A, Marmel F, Bair J, Kraus N. 2011. What subcortical–cortical relationships tell us about processing speech in noise. *Eur J Neurosci*. 33(3):549–557.

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. 2002. The processing of temporal pitch and melody information in auditory cortex. *Neuron*. 36(4):767–776.

Penagos H, Melcher JR, Oxenham AJ. 2004. A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *J Neurosci*. 24(30):6810–6815.

Presacco A, Simon JZ, Anderson S. 2016. Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *J Neurophysiol*. 116(5):2346–2355.

Price CN, Bidelman GM. 2021. Attention reinforces human corticofugal system to aid speech perception in noise. *Neuroimage*. 235:118014.

Puschmann S, Uppenkamp S, Kollmeier B, Thiel CM. 2010. Dichotic pitch activates pitch processing centre in Heschl's gyrus. *Neuroimage*. 49(2):1641–1649.

Ranieri F, Podda MV, Riccardi E, Frisullo G, Dileone M, Profice P, Pilato F, Di Lazzaro V, Grassi C. 2012. Modulation of LTP at rat hippocampal CA3-CA1 synapses by direct current stimulation. *J Neurophysiol*. 107(7):1868–1880.

Reato D, Rahman A, Bikson M, Parra LC. 2010. Low-intensity electrical stimulation affects network dynamics by modulating population rate and spike timing. *J Neurosci*. 30(45):15067–15079.

Ross B, Herdman AT, Pantev C. 2005. Right hemispheric laterality of human 40 Hz auditory steady-state responses. *Cereb Cortex*. 15(12):2029–2039.

Ross B, Tremblay KL, Alain C. 2020. Simultaneous EEG and MEG recordings reveal vocal pitch elicited cortical gamma oscillations in young and older adults. *Neuroimage*. 204:116253.

Russo NM, Skoe E, Trommer B, Nicol T, Zecker S, Bradlow A, Kraus N. 2008. Deficient brainstem encoding of pitch in children with autism spectrum disorders. *Clin Neurophysiol*. 119(8):1720–1731.

Salehinejad M, Kuo M, Nitsche M. 2019. The impact of chronotypes and time of the day on tDCS-induced motor cortex plasticity and cortical excitability. *Brain Stimul*. 12(2):421.

Schochat E, Rocha-Muniz CN, Filippini R. 2017. Understanding auditory processing disorder through the FFR. In: Kraus N, Anderson S, White-Schwoch T, Fay RR, Popper AN, editors. *The frequency-following response: a window into human communication*. Cham: Springer, pp. 225–250.

Shah B, Nguyen TT, Madhavan S. 2013. Polarity independent effects of cerebellar tDCS on short term ankle visuomotor learning. *Brain Stimul*. 6(6):966–968.

Skoe E, Kraus N. 2010. Auditory brainstem response to complex sounds: a tutorial. *Ear Hear*. 31(3):302.

Skoe E, Chandrasekaran B, Spitzer ER, Wong PC, Kraus N. 2014. Human brainstem plasticity: the interaction of stimulus probability and auditory learning. *Neurobiol Learn Mem*. 109:82–93.

Smalt CJ, Krishnan A, Bidelman GM, Ananthakrishnan S, Gandour JT. 2012. Distortion products and their influence on representation of pitch-relevant information in the human brainstem for unresolved harmonic complex tones. *Hear Res*. 292:26–34.

Song JH, Skoe E, Banai K, Kraus N. 2011. Perception of speech in noise: neural correlates. *J Cogn Neurosci*. 23(9):2268–2279.

Stagg CJ, Best JG, Stephenson MC, O'Shea J, Wylezinska M, Kincses ZT, Morris PG, Matthews PM, Johansen-Berg H. 2009. Polarity-sensitive modulation of cortical neurotransmitters by transcranial stimulation. *J Neurosci*. 29(16): 5202–5206.

Strait DL, Kraus N, Skoe E, Ashley R. 2009. Musical experience and neural efficiency: effects of training on subcortical processing of vocal expressions of emotion. *Eur J Neurosci*. 29: 661–668.

Suga N. 2020. Plasticity of the adult auditory system based on corticocortical and corticofugal modulations. *Neurosci Biobehav Rev*. 113:461–478.

Suga N, Gao E, Zhang Y, Ma X, Olsen JF. 2000. The corticofugal system for hearing: recent progress. *Proc Natl Acad Sci*. 97(22):11807–11814.

Schwenzer M, Mathiak K. 2011. Numeric aspects in pitch identification: an fMRI study. *BMC Neurosci*. 12(1):1–9.

Unal G, Bikson M. 2018. Transcranial Direct Current Stimulation (tDCS). In: Krames E, Peckham PH, Rezai AR, editors. *Neuromodulation: comprehensive textbook of principles, technologies, and therapies*. 2nd ed. London (United Kingdom): Academic Press, pp. 1589–1610.

Vanvooren S, Hofmann M, Poelmans H, Ghesquière P, Wouters J. 2015. Theta, beta and gamma rate modulations in the developing auditory system. *Hear Res*. 327:153–162.

White-Schwoch T, Carr KW, Thompson EC, Anderson S, Nicol T, Bradlow AR, Zecker SG, Kraus N. 2015. Auditory processing in noise: a preschool biomarker for literacy. *PLoS Biol*. 13(7):e1002196.

Willems RM, Van der Haegen L, Fisher SE, Francks C. 2014. On the other hand: including left-handers in cognitive neuroscience and neurogenetics. *Nat Rev Neurosci*. 15(3):193–201.

Wong PC, Skoe E, Russo NM, Dees T, Kraus N. 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci*. 10(4):420–422.

Wouters J, McDermott HJ, Francart T. 2015. Sound coding in cochlear implants: from electric pulses to hearing. *IEEE Signal Processing Magazine*. 32(2):67–80.

Yan J, Suga N. 1999. Corticofugal amplification of facilitative auditory responses of subcortical combination-sensitive neurons in the mustached bat. *J Neurophysiol*. 81(2):817–824.

Zaehle T, Beretta M, Jancke L, Herrmann CS, Sandmann P. 2011. Excitability changes induced in the human auditory cortex by transcranial direct current stimulation: direct electrophysiological evidence. *Exp Brain Res*. 215: 135–140.

Zatorre RJ, Belin P. 2001. Spectral and temporal processing in human auditory cortex. *Cereb Cortex*. 11(10): 946–953.

Zatorre RJ, Belin P, Penhune VB. 2002. Structure and function of auditory cortex: music and speech. *Trends Cogn Sci*. 6(1):37–46.

Zatorre RJ, Evans AC, Meyer E, Gjedde A. 1992. Lateralization of phonetic and pitch discrimination in speech processing. *Science*. 256(5058):846–849.

Zhang Y, Suga N. 1997. Corticofugal amplification of subcortical responses to single tone stimuli in the mustached bat. *J Neurophysiol*. 78(6):3489–3492.